EMPIRICAL BAYES MINIMAX ESTIMATORS OF MATRIX NORMAL MEANS IN THE ONE AND TWO-SAMPLE PROBLEMS

BY

GWOWEN SHIEH

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1990

Copyright 1990

by

Gwowen Shieh

To my wife Show-Li

ACKNOWLEDGEMENTS

I would like to thank Dr. Malay Ghosh, my dissertation advisor, for his encouragement, patience, and invaluable support. Also, I would like to thank Dr. Mark C. K. Yang, Dr. Beverly L. Brechner, Dr. Pejaver V. Rao, and Dr. Rocco Ballerini for serving on my committee. In addition, I would like to express my appreciation to Dr. Ralph G. O'Brien for his encouragement and kindness of letting me use his computer facilities.

Lastly, I would like to thank my parents and parents-in-law for their understanding, encouragement, and continual support.

TABLES OF CONTENTS

ACKNOWLEDGEMENTS iv

LIST OF TABLES vii

ABSTRACT ix

CHAPTERS

1 INTRODUCTION 1

1.1 Literature Review 1
1.2 Subject of Research 5
1.3 Some Tools 7

2 EMPIRICAL BAYES MINIMAX ESTIMATORS OF MATRIX NORMAL MEANS FOR ARBITRARY QUADRATIC LOSS AND KNOWN COVARIANCE MATRIX IN THE ONE SAMPLE PROBLEM 9

2.1 Introduction 9
2.2 EB Minimax Estimators Under The L_Q Loss 12
2.3 Improvement On $(\underline{I}_p - r(\underline{Y}, \underline{S})\underline{V}\underline{S}^{-1})\underline{X}$ 24
2.4 Simulation Results 26

3 EMPIRICAL BAYES MINIMAX ESTIMATORS OF MATRIX NORMAL MEANS FOR ARBITRARY QUADRATIC LOSS AND UNKNOWN COVARIANCE MATRIX IN THE ONE-SAMPLE PROBLEM 33

3.1 Introduction 33
3.2 EB Minimax Estimators Under The L_{Q^*} Loss and $\underline{V} = v\underline{G}$, Only v Unknown 38
3.3 EB Minimax Estimators Under The L_{Q^*} Loss and $\underline{V} = \underline{V}_d$ an Unknown Diagonal Matrix 45
3.4 EB Minimax Estimators Under The L_{Q^*} Loss and $\underline{V}$ Is Completely Unknown 58
3.5 Simulation Results 69

4 EMPIRICAL BAYES MINIMAX ESTIMATORS OF MATRIX NORMAL MEANS FOR ARBITRARY QUADRATIC LOSS AND KNOWN COVARIANCE MATRIX IN THE TWO SAMPLE PROBLEM 84

4.1 Introduction 84
4.2 EB Minimax Estimators Under The L_Q Loss 88

4.3 Improvement On PTE 92

5 EMPIRICAL BAYES MINIMAX ESTIMATORS OF MATRIX NORMAL MEANS FOR ARBITRARY QUADRATIC LOSS AND UNKNOWN COVARIANCE MATRIX IN THE TWO-SAMPLE PROBLEM 97

5.1 Introduction 97
5.2 EB Minimax Estimators Under The L_{Q^*} Loss 99
5.3 Improvement On PTE 103

6 SUMMARY AND FUTURE RESEARCH 108

6.1 Summary 108
6.2 Future Research 109

BIBLIOGRAPHY 110

BIOGRAPHICAL SKETCH 113

LIST OF TABLES

2 1 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_1 = \underline{0}$, $\underline{V}$ known) 29

2 2 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_1 = \underline{0}$, $\underline{V}$ known) 30

2 3 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_1 = (10, 35)^T$, $\underline{V}$ known) 31

2 4 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_1 = (6, 10, 50)^T$, $\underline{V}$ known) 32

3 1 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_1 = \underline{0}$, $\underline{V}$ known up to a positive constant) 72

3 2 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_1 = \underline{0}$, $\underline{V}$ known up to a positive constant) 73

3 3 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_1 = (10, 35)^T$, $\underline{V}$ known up to a positive constant) 74

3 4 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_1 = (6, 10, 50)^T$, $\underline{V}$ known up to a positive constant) 75

3 5 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_1 = \underline{0}$, $\underline{V}$ unknown diagonal matrix) 76

3 6 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_1 = \underline{0}$, $\underline{V}$ unknown diagonal matrix) 77

3 7 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_1 = (10, 35)^T$, $\underline{V}$ unknown diagonal matrix) 78

3 8 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_1 = (6, 10, 50)^T$, $\underline{V}$ unknown diagonal matrix) 79

3 9 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_1 = \underline{0}$, $\underline{V}$ completely unknown) 80

3 10 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_1 = \underline{0}$, $\underline{V}$ completely unknown) 81

3 11 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_1 = (10, 35)^T$, $\underline{V}$ completely unknown) 82

3 12 THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_1 = (6, 10, 50)^T$, $\underline{V}$ completely unknown) 83

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

EMPIRICAL BAYES MINIMAX ESTIMATORS OF MATRIX NORMAL MEANS IN THE ONE- AND TWO-SAMPLE PROBLEMS

By

GWOWEN SHIEH

December, 1990

Chairman: Dr. Malay Ghosh
Major Department: Statistics

This dissertation is concerned with the estimation of a matrix normal means in one- and two-sample problems. In both cases the empirical Bayes method of estimation is used. We consider situations of both known and unknown covariance matrices when the former is not necessarily an identity matrix, while the latter has three possible structures. The performance of estimator is evaluated in terms of risks under the quadratic loss with different weight matrices. Those estimators that are considered do not depend on the weight matrix. In the one-sample problem, the proposed EB estimators are uniformly better than the sample mean, a well-known minimax estimator, and therefore are minimax. In the two-sample problem, the first sample mean is a minimax estimator of the corresponding population normal means and is dominated by proposed EB estimators. In this case, they are weighted averages of the first sample mean and pooled sample mean, and incorporate in a very natural way that the degree of evidence for or against the hypothesis of the equality of two population normal means. Then the preliminary test estimator is improved by the modified EB estimator which equals the pooled sample mean when the null hypothesis is accepted at a desired level of significance and equals the EB estimator when the opposite is the case.

CHAPTER 1
INTRODUCTION

1.1 Literature Review

Stein (1956), in his seminal paper, discovered the surprising result that the sample mean vector $\underline{X}$ of a K (> 2) dimensional normal distribution $N_K(\underline{\theta}, I_K)$ is an inadmissible estimator of its mean $\underline{\theta}$ under the squared error loss $L(\underline{\theta}, \underline{a}) = (\underline{\theta}-\underline{a})^T(\underline{\theta}-\underline{a})$, where I_K is the identity matrix of order K. An explicit estimator dominating the sample mean was provided by James and Stein (1961) under the same setup. Since then, a large body of research has been devoted to finding uniformly better estimators under different assumptions about the covariance structure $\underline{M}$ of $\underline{X}$ and under different quadratic losses $L_1(\underline{\theta}, \underline{a}) = (\underline{\theta}-\underline{a})^T Q(\underline{\theta}-\underline{a})$ or $L_2(\underline{\theta}, \underline{a}) = (\underline{\theta}-\underline{a})^T Q(\underline{\theta}-\underline{a})/\mathrm{tr}(Q\underline{M})$, where Q is a known K×K symmetric positive definite (p.d.) matrix. Since the sample mean is a constant risk minimax estimator of $\underline{\theta}$ in any arbitrary dimension under the L_1 loss for a known $\underline{M}$ or under the L_2 loss for known or unknown $\underline{M}$, any estimator dominating the sample mean is also a minimax estimator of $\underline{\theta}$. Much of the earlier research on this topic has concentrated on finding a general class of minimax estimators in cases where $\underline{M}$ is known and $Q = I_K$, or $\underline{M}$ is known up to a positive multiple and $Q = I_K$ or when $Q = \underline{M}^{-1}$ and $\underline{M}$ is known or unknown. This research initiated by Baranchik (1970), and was culminated eventually in the work of Efron and Morris (1976a). For the more general situation, Berger and Bock (1976a, 1976b) and Shinozaki (1977) have found minimax estimators under the assumption that $\underline{M}$ is an unknown diagonal matrix. The case of a completely unknown $\underline{M}$ is more complex, and thus has been more resistant to solution. Berger et al. (1977), Gleser (1979, 1986a, 1986b) and Berger and Haff (1983) have succeeded in developing estimators which dominate $\underline{X}$ in risks when no restrictions are placed on $\underline{M}$ and Q. However, these minimax estimators involve the weight matrix Q which appears in the loss, and are quite sensitive to the choice of Q. Brown (1975) and Shinozaki (1980) have provided sufficient conditions under which a class of estimators not depending on Q dominates the sample mean simultaneously under several diagonal Q's and the identity covariance matrix. However, the class of Q's for which the dominance holds turns out to be very narrow. For the case that covariance matrix is known upto a constant multiple and Q is an arbitrary p.d. matrix, Ghosh (1989) has developed a wide class of minimax estimators not depending on Q. This class of estimators turns out to be a generalization of the class of minimax estimators corresponding to the case where the weight matrix is strictly an identity matrix. Slightly generalized version of the result of Ghosh (1989) is stated below without proof.

Theorem 1.1.1. Let $\underline{X} \sim N_K(\underline{\theta}, \sigma^2 I_K)$ where σ^2 (> 0) is an unknown constant, but an estimator $\widehat{\sigma^2}$, independent of $\underline{X}$ and distributed as $\sigma^2(n+2)^{-1}\chi^2_n$, is available. Assume that $tr(Q) > 2ch_1(Q)$, where $ch_1(Q)$ denotes the largest eigen value of Q, and $\tau(F)$ is a scalar function differentiable in F with the derivative $\tau'(F)$ and $F = (\underline{X}^T\underline{X})/\widehat{\sigma^2}$. Suppose

(i) $0 < \tau(F) < 2[tr(Q)/ch_1(Q)-2]$,

(ii) $F^{(K-2)/2}\tau(F)/(2-\tau(F))^{1+2(K-2)/(n+2)}$ is increasing in F.

Then $(1-\tau(F)/F)\underline{X}$ dominates $\underline{X}$ under the quadratic loss $L(\underline{\theta}, \underline{a}) = (\underline{\theta}-\underline{a})^TQ(\underline{\theta}-\underline{a})$ or $L(\underline{\theta}, \underline{a}) = (\underline{\theta}-\underline{a})^TQ(\underline{\theta}-\underline{a})/tr(\sigma^2Q)$.

Stein's phenomenon can be extended to the problem of estimating a matrix of normal means. Suppose $\underline{X}_1, \ldots, \underline{X}_K$ are K independent p-dimensional column vectors (K > p+1) such that $\underline{X}_i$ is normally distributed with mean vector $\underline{\theta}_i$ and the identity covariance matrix I_p. In this case, the sample mean matrix $X = (\underline{X}_1, \ldots, \underline{X}_K)$ is a minimax estimator of

$\underline{\theta} = (\underline{\theta}_1, \ldots, \underline{\theta}_K)$ under the squared error loss $L(\underline{\theta}, \underline{a}) = \mathrm{tr}\{(\underline{\theta}-\underline{a})^T(\underline{\theta}-\underline{a})\}$ Efron and Morris (1972b, 1976b), Stein (1974), Haff (1977, 1979), Dey, Ghosh and Srinivasan (1985) and Zheng (1986) have provided estimators that are uniformly better than $\underline{X}$ thereby proving the inadmissibility of $\underline{X}$. However, these results are restricted to the identity covariance matrix and squared error loss. For the more general situation of quadratic loss and covariance matrix $\underline{V}$ of $\underline{X}_i$'s, the results thus far obtained are very incomplete. Efron and Morris (1972b) have considered the problem when the p.d. covariance matrix $\underline{V}$ is completely known (but not necessarily $\underline{I}_p$), known up to a positive scalar constant, unknown diagonal matrix and completely unknown matrix. Under the special loss function $L(\underline{\theta}, \underline{a}) = \mathrm{tr}\{(\underline{\theta}-\underline{a})^T\underline{V}^{-1}(\underline{\theta}-\underline{a})\}$, these authors proposed the matrix extension form James-Stein estimator $(\underline{I}_p-(K-p-1)\underline{V}\,\underline{S}^{-1})\underline{X}$ when $\underline{V}$ is known where $\underline{S} = \underline{X}^T\underline{X}$. The estimator $(\underline{I}_p-(K-p-1)\hat{\underline{V}}\,\underline{S}^{-1})\underline{X}$ was proposed when $\underline{V}$ is unknown and $\hat{\underline{V}}$, independent of $\underline{X}$, is an estimate of $\underline{V}$

For estimating the vector normal mean, Efron and Morris (1972a, 1973a) motivated the James-Stein estimator from an empirical Bayes (EB) point of view They continued the effort to motivate the estimators described in the previous paragraph as EB estimators For $\underline{Q} = \underline{V} = \underline{I}_p$, more general EB estimators were provided by Haff (1979) and Dey, Ghosh and Shrinivasan (1985)

One of the objectives of this dissertation is to provide EB minimax estimators of matrix normal means $\underline{\theta}$ for the case where $\underline{Q}$ is arbitrary p.d and $\underline{V}$ is p.d. (known or unknown) All the different structures of the covariance matrix $\underline{V}$ given in Efron and Morris (1972b) are covered. The loss function being considered is $L_{\underline{Q}}(\underline{\theta}, \underline{a}) = \mathrm{tr}\{(\underline{\theta}-\underline{a})^T\underline{Q}(\underline{\theta}-\underline{a})\}$ when $\underline{V}$ is known or $L_{\underline{Q}^*}(\underline{\theta}, \underline{a}) = \mathrm{tr}\{(\underline{\theta}-\underline{a})^T\underline{Q}^*(\underline{\theta}-\underline{a})\}$ when $\underline{V}$ is known or unknown, where $\underline{Q}^* = \underline{V}^{-1/2}\underline{Q}\,\underline{V}^{-1/2}$ and $\underline{V} = \underline{V}^{1/2}\underline{V}^{1/2}$

The principal idea behind an EB approach is to "borrow strength from the ensemble" to construct estimators In Efron and Morris (1973b), the EB idea was generalized to combine possibly related estimation problems when there are two sets of parameters to estimate, and decide whether the James-Stein estimator should be applied separately to the two sets or to the combined set. These two groups of parameters are suspected of having same zero prior mean and different prior variances. Efron and Morris (1973b) imposed a distribution on the ratio of the prior variances and derived a class of compromise estimators which is superior than either alternatives in terms of Bayes risks under the squared error loss. Berger and Dey (1983) have also considered the problem of combining groups when simultaneously estimating normal means. Their situation is substantially different from the one in Efron and Morris (1973b) since no symmetries and EB relationships were assumed to exist. These authors applied two types of estimators and found that the so called 'combined' estimator had smaller Bayes risk under the quadratic loss. Another type of grouping problem can be found in Ghosh and Sinha (1988). They considered the situation when there were two groups of observations but interest lay in estimating the vector normal means of the first sample. The question is whether the pooled sample mean or the first sample mean should be used to estimate the first group of normal means when the two groups of normal means are suspected to be equal. Both EB and hierarchical Bayes (HB) estimators were proposed and shown to dominate the first sample mean under the quadratic loss. Unlike Berger and Dey (1983), these estimators did not involve the weight matrix Q. Further, modified EB estimators were proposed and sufficient conditions under which they dominated the preliminary test estimators (PTE) are given. The PTE uses the pooled sample mean when the null hypothesis of the equality of the two population means is accepted at a desired level of significance and uses the first sample mean when the opposite is the case.

The second objective of this research generalizes the problem described above into the estimation of a matrix normal means when two matrix sample means are available and they are suspected to have same population matrix normal means. The proposed EB and modified EB estimators are evaluated in terms of risks under the quadratic loss $L_Q(\theta, a) =$

$\mathrm{tr}[(\underline{\theta}-\underline{a})^T Q(\underline{\theta}-\underline{a})]$ when the covariance matrix is known or $L_{Q^*}(\underline{\theta}, \underline{a}) = \mathrm{tr}[(\underline{\theta}-\underline{a})^T Q^*(\underline{\theta}-\underline{a})]$ when the covariance matrix is unknown.

1.2 Subject of Research

The present research focuses on developing the EB minimax estimators of matrix normal means in one- and two-sample problems. Although, it is known from the previous literature that Stein's phenomenon exists in the problem of estimation of matrix normal means, our study addresses the problem in greater generality situations. Three goals have been set and achieved. First, a general covariance matrix $\underline{V}$ (known or unknown) is considered. Second, we consider a general quadratic loss $L_Q(\underline{\theta}, \underline{a}) = \mathrm{tr}\{(\underline{\theta}-\underline{a})^T Q(\underline{\theta}-\underline{a})]$ or $L_{Q^*}(\underline{\theta}, \underline{a}) = \mathrm{tr}\{(\underline{\theta}-\underline{a})^T Q^*(\underline{\theta}-\underline{a})\}$, unlike the previous work on matrix normal means estimation where $Q = I_p$. Finally, the proposed estimators dominating the matrix sample mean do not depend on the weight matrix Q, and hence they are robust against the misspecification of Q.

Both Chapter 2 and Chapter 3 are devoted to the estimation of matrix normal means in the one-sample problem. In Chapter 2, we assume a completely known arbitrary covariance matrix, while in Chapter 3 we assume the covariance matrix to be unknown. Three cases are considered in Chapter 3. First, the covariance matrix is assumed to be known up to a positive constant multiple, second, an unknown diagonal matrix, and third, a completely unknown matrix. It is assumed that in each case, there is an estimate available to estimate the unknown covariance matrix and the estimate is independent of the observed sample mean. These are discussed in Sections 3.2-3.4, respectively.

In Chapter 4 and Chapter 5, we consider the estimation of matrix normal means in the two-sample problem. This particular study of combining related estimation problems stems directly from Ghosh and Sinha (1988). The analyses of the first parts in both Chapter 4 and Chapter 5 are very much parallel to the analyses in Chapter 2 and Chapter 3

In an EB framework, we propose a general form of EB estimator $\hat{\underline{\theta}}^{EB}$ ($\hat{\underline{\theta}}_1^{EB}$ in Chapter 4 and Chapter 5) which bears a strong resemblance to Baranchik's (1970). Actually they are exactly the same when the dimension of the matrix normal means reduce from p×K to 1×K. Two classes of EB estimators, $\hat{\underline{\theta}}^{EB(1)}$ and $\hat{\underline{\theta}}^{EB(2)}$ ($\hat{\underline{\theta}}_1^{EB(1)}$ and $\hat{\underline{\theta}}_1^{EB(2)}$), are derived and are shown to be uniformly better than the sample mean in one-sample problem (or uniformly better than the first sample mean in two-sample problem) under the quadratic loss $L_Q(\underline{\theta}, \underline{a}) = \mathrm{tr}\{(\underline{\theta}-\underline{a})^T Q(\underline{\theta}-\underline{a})\}$ or $L_{Q^*}(\underline{\theta}, \underline{a}) = \mathrm{tr}\{(\underline{\theta}-\underline{a})^T Q^*(\underline{\theta}-\underline{a})\}$. We have also shown under certain conditions $\hat{\underline{\theta}}^{EB(1)}$ ($\hat{\underline{\theta}}_1^{EB(1)}$) is improved by $\hat{\underline{\theta}}^{EB(3)}$ ($\hat{\underline{\theta}}_1^{EB(3)}$). Since the matrix extension form of James-Stein estimator belongs to the first class, our result shows its inadmissibility for the estimation of matrix normal means while no estimator (except its own positive-part version) has been found to dominate the James-Stein estimator for the estimation of vector normal means.

In Sections 4.3 and 5.3, the problem of how to reach a compromise between the first sample mean and pooled sample mean is investigated. In problems of this type, what is normally sought is a compromise estimator which leans more towards the pooled sample mean when the hypothesis of the equality of the two population means is accepted, and towards the first sample mean when such a hypothesis is rejected. One classical way to achieve this compromise is to use the PTE, although the procedure suffers from the drawback that there is no general way of incorporating the degree of evidence for or against the null hypothesis in order to arrive at the estimator. For instance, if the observed statistic values are very close to a critical value at a chosen level of significance, then there may be no persuasive reason for choosing one estimator in preference to the other. However, the EB minimax estimators ($\hat{\underline{\theta}}_1^{EB(1)}$, $\hat{\underline{\theta}}_1^{EB(2)}$ and $\hat{\underline{\theta}}_1^{EB(3)}$), proposed in Sections 4.2 and 5.2 incorporate the degrees of evidence for or against the null hypothesis in a very natural way. Further, their modified versions ($\hat{\underline{\theta}}_1^{MEB(1)}$, $\hat{\underline{\theta}}_1^{MEB(2)}$ and $\hat{\underline{\theta}}_1^{MEB(3)}$) are shown to dominate the corresponding PTE in terms of risks. Such a modified EB minimax estimator continues to be a weighted average of the first sample mean and the pooled sample mean It equals the pooled sample mean when the hypothesis is accepted at a desired level of significance, and equals the EB minimax estimator when the opposite is the case.

Numerical studies are undertaken in Section 2.4 and Section 3.5 to illustrate the applicability of the estimators proposed in Section 2.2-2 3 and Section 3.2 3 4, respectively Finally, a summary of the present work and some suggestions for future study are given in Chapter 6

1.3 Some Tools

In deriving the minimax estimators for the vector normal means the techniques which have been found to be most useful are Stein's (1974) integration-by-parts identity for the normal distribution, which permits construction of unbiased estimators of risk difference for competing estimators of $\underline{\theta}$ and Wishart identity due to Haff (1982) A scalar version of the Wishart identity, usually referred to as the chi-square identity, appeared in Efron and Morris (1976a) These identities are now given below

Lemma 1.3.1 (Stein's identity) Suppose $\underline{X}$, a p×1 column vector, has a multivariate normal distribution $N_p(\underline{\theta}, \sigma^2 \underline{I}_p)$ and $h(\underline{X})$ is differentiable with respect to X_i, i = 1, ..., p Then

$$E_{\underline{\theta},\sigma^2}\left[\frac{\partial h(\underline{X})}{\partial \underline{X}}\right] = E_{\underline{\theta},\sigma^2}\left[\frac{h(\underline{X})}{\sigma^2}(\underline{X}-\underline{\theta})\right], \tag{1 3 1}$$

where $\frac{\partial h(\underline{X})}{\partial \underline{X}}$ is a column vector with ith element equal to $\frac{\partial h(\underline{X})}{\partial X_i}$, i = 1 ... p

Lemma 1.3.2 (Chi-square identity) Suppose $\hat{v}$ is distributed as $v(n+2)^{-1}\chi^2_n$, where v (> 0) is a constant and χ^2_n is the chi-square distribution with n degrees of freedom. Let $w(\hat{v})$ be differentiable in $\hat{v}$ and $w'(\hat{v})$ denote the derivative of $w(\hat{v})$. Then

$$E_v[w(\hat{v})\hat{v}v^{-1}] = E_v[\frac{n}{(n+2)}w(\hat{v})+\frac{2}{(n+2)}w'(\hat{v})\hat{v}] \qquad (1.3.2)$$

Lemma 1.3.3 (Wishart identity) Assume that $\underline{S} = ((s_{ij}))$, a p×p matrix, has a nonsingular Wishart distribution with matrix $\underline{\Sigma}$ and K degrees of freedom, i.e., $\underline{S} \sim W(p, K, \underline{\Sigma})$, where $K > p+1$. Let $\underline{D}^S = ((d^S_{ij}F))$ and $d^S_{ij} = 1/2(1+\delta_{ij})\partial/\partial s_{ij}$, δ_{ij} is the Kronecker delta, $i, j = 1, \ldots, p$. Then for a suitable p×p matrix $\underline{W} = \underline{W}(\underline{S})$, we have the Wishart identity

$$E_{\underline{\Sigma}}[\underline{\Sigma}^{-1}\underline{W}(\underline{S})] = E_{\underline{\Sigma}}[2\underline{D}^S\underline{W}(\underline{S})+(K-p-1)\underline{S}^{-1}\underline{W}(\underline{S})] \qquad (1.3.3)$$

Note that the scalar (trace) version of this identity was derived independently by Stein (1974) and Haff (1979).

CHAPTER 2
EMPIRICAL BAYES MINIMAX ESTIMATORS OF MATRIX NORMAL MEANS FOR ARBITRARY QUADRATIC LOSS AND KNOWN COVARIANCE MATRIX IN THE ONE-SAMPLE PROBLEM

2.1 Introduction

The problem of estimation of matrix normal means can be stated as follows. Let $\underline{X}_1, \ldots, \underline{X}_K$ be K independent p-dimensional column vectors, $K > p+1$, with $\underline{X}_i$ having conditional distribution

$$\underline{X}_i \mid \underline{\theta}_i \sim N_p(\underline{\theta}_i, \underline{V}), \tag{2.1.1}$$

where $\underline{\theta}_i$'s are unknown, but $\underline{V}$ is a known $p \times p$ symmetric p.d. matrix. Later in Chapter 3, we will consider an unknown $\underline{V}$. Let $\underline{\theta} = (\underline{\theta}_1, \ldots, \underline{\theta}_K)$, $\underline{X} = (\underline{X}_1, \ldots, \underline{X}_K)$. The problem is to estimate $\underline{\theta}$. For example, in a balanced one-way multivariate analysis of variance (MANOVA) model

$$\underline{X}_{ij} = \underline{\theta}_i + \underline{e}_{ij} \; (i = 1, \ldots, K, \; j = 1, \ldots, m), \tag{2.1.2}$$

where $\underline{X}_{ij}$ is a $p \times 1$ vector of responses of the jth individual receiving the ith treatment, and $\underline{\theta}_i$ is the $p \times 1$ vector of ith treatment effects. The vectors $\underline{e}_{ij}$'s are independently distributed as $N_p(\underline{0}, \underline{V}_0)$, where $\underline{V}_0$ is a known $p \times p$ symmetric p.d. matrix. The minimal sufficient statistic for $\underline{\theta} = (\underline{\theta}_1, \ldots, \underline{\theta}_K)$ is $\underline{X} = (\underline{X}_1, \ldots, \underline{X}_K)$, where $\underline{X}_i = m^{-1}\sum_{j=1}^{m} \underline{X}_{ij}$. Note that the $\underline{X}_i$'s are independently distributed as $N_p(\underline{\theta}_i, \underline{V})$, where $\underline{V} = m^{-1}\underline{V}_0$ and $K > p+1$.

In a Bayesian framework, (2.1.1) is interpreted as the conditional distribution of $\underline{X}_i$ given $\underline{\theta}_i$. Suppose also that $\underline{\theta}_i$'s have independent distributions

$$\underline{\theta}_i \sim N_p(\underline{0}, \underline{A}), \tag{2.1.3}$$

where $\underline{\Delta}$ is a p.d. matrix. Then the posterior distribution of $\underline{\theta}_i$ given $\underline{X}_i = \underline{x}_i$ is

$$\underline{\theta}_i \mid \underline{X}_i = \underline{x}_i \sim N_p((\underline{I}_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{x}_i,\ (\underline{I}_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{V}), \tag{2.1.4}$$

where $\underline{I}_p$ is the identity matrix of order p and $\underline{\Sigma} = (\underline{V} + \underline{\Delta})$. Under the quadratic loss

$$L_Q(\underline{\theta}, \underline{a}) = \sum_{i=1}^{K} (\underline{\theta}_i - \underline{a}_i)^T \underline{Q} (\underline{\theta}_i - \underline{a}_i) \tag{2.1.5}$$

where $\underline{Q}$ is p.d., the Bayes estimator of $\underline{\theta}$ is given by

$$\hat{\underline{\theta}}^{B}(\underline{X}) = (\underline{I}_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{X}. \tag{2.1.6}$$

In the EB scenario, $\underline{\Delta}$ and $\underline{\Sigma}$ are unknown, and are estimated from the marginal distribution of $\underline{X}$. Note that the marginal distributions of $\underline{X}_i$'s are mutually independent with

$$\underline{X}_i \sim N_p(\underline{0}, \underline{\Sigma}), \tag{2.1.7}$$

and $\underline{S} = \sum_{i=1}^{K} \underline{X}_i \underline{X}_i^T = \underline{X}\,\underline{X}^T$ is complete sufficient for $\underline{\Sigma}$, where $\underline{S} \sim$ Wishart $W(p, K, \underline{\Sigma})$. A natural candidate for the choice of $\underline{\Sigma}^{-1}$ is $(K-p-1)\underline{S}^{-1}$, the best multiple of $\underline{S}^{-1}$. More generally, one can consider the class of EB estimators

$$\hat{\underline{\theta}}^{EB}(\underline{X}) = (\underline{I}_p - \underline{V}\,\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S}))\underline{X} \tag{2.1.8}$$

of $\underline{\theta}$, where $\underline{\tau}(\underline{V}, \underline{S})$ is a p×p symmetric matrix. Since $\underline{V}$ is known, the appearance of $\underline{V}$ in $\underline{\tau}(\underline{V}, \underline{S})$ is superfluous. But we prefer to continue with this notation, since later for unknown $\underline{V}$, we will use $\underline{\tau}(\hat{\underline{V}}, \underline{S})$, where $\hat{\underline{V}}$ estimates $\underline{V}$.

Estimation of the matrix normal means was also considered in Efron and Morris (1972b, 1976b), Stein (1974), Haff (1977, 1979), Dey, Ghosh and Srinivasan (1985) and Zheng (1986) among others. Efron and Morris (1972b) showed that $(\underline{I}_p - (K-p-1)\underline{S}^{-1})\underline{X}$ dominated $\underline{X}$ uniformly in $\underline{\theta}$ under the loss (2.1.5) with $\underline{Q} = \underline{I}_p$, i.e., the squared error loss, and $\underline{V} = I_p$. Later, Stein (1974) and Efron and Morris (1976b) proposed estimators of the form $(I_p - a\underline{S}^{-1} - b(\mathrm{tr}\,\underline{S})^{-1}I_p)\underline{X}$ for $\underline{\theta}$ which dominate $\underline{X}$ for certain choices of the constants a and b, once again for $\underline{V} = I_p$ and the squared error loss. Since $\underline{X}$ is a minimax estimator of $\underline{\theta}$ under the quadratic loss (2.1.5) for every p.d. Q, any estimator dominating $\underline{X}$ is also minimax. Zheng (1986) provided a general class of minimax estimators of $\underline{\theta}$ which included as its members the ones mentioned earlier for $Q = \underline{V} = I_p$. In particular, his Theorem 2 (page 75) generalized Baranchik's (1970) results on minimax estimation of the vector of normal means.

It will be shown in Section 2.2 that EB estimators of the form (2.1.8) dominate $\underline{X}$ under the loss (2.1.5). Thus, the results in this section generalize the previous results in two directions. First, $\underline{V}$ is not necessarily I_p. Second, the loss is the more general (2.1.5) with an arbitrary Q instead of $Q = I_p$. We shall see that the problem of finding estimators which dominate $\underline{X}$ under the loss (2.1.5) for every p.d. Q amounts to the problem of finding solutions in differential inequalities of a certain type. As a special case (Theorem 2.2.3), it is shown that Zheng's Theorem 2 continues to hold under the more general loss (2.1.5) when $\underline{V}$ is not necessarily I_p. However, other minimax estimators will be proposed where conditions required in Theorem 2 of Zheng (1986) are not met. Some of these estimators are different from the ones proposed in Haff (1979) and Dey, Ghosh and Srinivasan (1985).

It will be shown in Section 2.2 (Example 2.2.1) that EB estimators of the form $(I_p - a\underline{V}\,\underline{S}^{-1})\underline{X}$ dominate $\underline{X}$ for every $0 < a < 2(K-p-1)$ under the loss (2.1.5). The best choice of a is $K-p-1$. Stein (1974) and Efron and Morris (1976b) provided estimators dominating $(I_p - (K-p-1)\underline{V}\,\underline{S}^{-1})\underline{X}$ when $\underline{V} = I_p$ and the loss was squared error. Under the same set up, Haff (1979) provided estimators dominating $(I_p - a\underline{S}^{-1})\underline{X}$ for $0 < a \leq K-p-1$. We provide in Section 2.3 several estimators which dominate $(I_p - r(\underline{V}, \underline{S})\underline{V}\,\underline{S}^{-1})\underline{X}$ for $0 < r(\underline{V}, \underline{S}) < K-2$. Again, the previous results are generalized in two directions, i.e., $\underline{V}$ is not necessarily I_p and the loss is (2.1.5) instead of squared error loss. Moreover, it will be encouraging to find that some of the estimators for $\underline{\Sigma}^{-1}$ proposed by Haff (1979) or Dey, Ghosh and Srinivasan (1985) can be used for $\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S})$ in (2.1.8) and those estimators dominate $(\underline{I}_p - a\underline{V}\,\underline{S}^{-1})\underline{X}$ for $0 < a < K-2$ under the loss (2.1.5).

In Section 2.4, we have provided the simulated risks of five different estimators of $\underline{\theta}$ under the loss (2.1.5) divided by $tr(\underline{Q}\,\underline{V})$ for different choice of p, K, $\underline{\theta}$, $\underline{V}$ and $\underline{Q}$. It turns out that the estimator $\widehat{\underline{\theta}}_{\alpha_0}(\underline{X}) = (\underline{I}_p - \underline{V}[(1-\alpha_0)[(K-p-1)\underline{S}^{-1} + ((p-1)/tr(\underline{S}))\underline{I}_p] + \alpha_0[(Kp-2)/tr(\underline{S})]\underline{I}_p])\underline{X}$ with $\alpha_0 = (K-p-1)/[K^2(p-1)]$ has a definite edge over its rival estimators when $\underline{\theta} = \underline{0}$. However, for $\underline{\theta}$ far away from $\underline{0}$, among the estimators dominating $\underline{X}$, there is no clear-cut advantage of any particular estimator over the others. Thus, on an overall basis, we recommend the use of $\widehat{\underline{\theta}}_{\alpha_0}(\underline{X})$ as an estimator of $\underline{\theta}$.

2.2 EB Minimax Estimators Under The L_Q Loss

We first prove a result which connects the risk of $\underline{X}$ with the risk of $\widehat{\underline{\theta}}^{EB}(\underline{X})$ defined in (2.1.8). Two important corollaries are derived concerning two special classes of EB estimators of the form (2.1.8). We denote by $E_{\underline{\theta}}$, the expectation conditional on $\underline{\theta}$, $\overline{E}$ the expectation over the marginal distribution of $\underline{\theta}$, and E the expectation over the joint distribution of $\underline{X}$ and $\underline{\theta}$. Also, let $R_Q(\underline{\theta}, \underline{g}(\underline{X})) = E_{\underline{\theta}}[L_Q(\underline{\theta}, \underline{g}(\underline{X}))]$, where L_Q is defined in (2.1.5).

Theorem 2.2.1 Under the L_Q loss given in (2.1.5), $\widehat{\underline{\theta}}^{EB}$ defined in (2.1.8) has risk

$$R_Q(\underline{\theta}, \widehat{\underline{\theta}}^{EB}) = R_Q(\underline{\theta}, \underline{X}) + E_{\underline{\theta}} tr\{[\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S})\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1} - 4\underline{D}^{\underline{S}}(\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1})$$

$$- 2(K-p-1)\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}]\underline{Q}_0\}, \qquad (2.2.1)$$

where $\underline{Q}_0 = \underline{V}\,\underline{Q}\,\underline{V}$ and $\underline{D}^{\underline{S}}$ is as defined in Lemma 1.3.3

Proof of Theorem 2.2.1 Let $\widehat{\underline{\theta}}_i^{B}(\underline{X}) = (I_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{X}_i$ and $\widehat{\underline{\theta}}_i^{EB}(\underline{X}) = (I_p - \underline{V}\,\underline{S}^{-1}\tau(\underline{V}, \underline{S}))\underline{X}_i$, $i = 1, \ldots, K$. Then,

$$EL_Q(\underline{\theta}, \widehat{\underline{\theta}}^{EB})$$

$$= E\{\sum_{i=1}^{K}(\underline{\theta}_i - \widehat{\underline{\theta}}_i^{EB})^T Q(\underline{\theta}_i - \widehat{\underline{\theta}}_i^{EB})\}$$

$$= E\{\sum_{i=1}^{K}(\underline{\theta}_i - \widehat{\underline{\theta}}_i^{B} + \widehat{\underline{\theta}}_i^{B} - \widehat{\underline{\theta}}_i^{EB})^T Q(\underline{\theta}_i - \widehat{\underline{\theta}}_i^{B} + \widehat{\underline{\theta}}_i^{B} - \widehat{\underline{\theta}}_i^{EB})\}$$

$$= EL_Q(\underline{\theta}, \widehat{\underline{\theta}}^{B}) + E\{\sum_{i=1}^{K}(\widehat{\underline{\theta}}_i^{B} - \widehat{\underline{\theta}}_i^{EB})^T Q(\widehat{\underline{\theta}}_i^{B} - \widehat{\underline{\theta}}_i^{EB})\}$$

$$= EL_Q(\underline{\theta}, \widehat{\underline{\theta}}^{B}) + E\mathrm{tr}[(\underline{S}^{-1}\tau(\underline{V}, \underline{S}) - \underline{\Sigma}^{-1})\underline{S}(\tau(\underline{V}, \underline{S})\underline{S}^{-1} - \underline{\Sigma}^{-1})Q_0] \tag{2.2.2}$$

Next, using the fact that $\underline{\theta}_i \mid \underline{X}_i$ has distribution (2.1.4), one has

$$EL_Q(\underline{\theta}, \widehat{\underline{\theta}}^{B}) = K\mathrm{tr}\{(I_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{V}\,Q\} \tag{2.2.3}$$

Also,

$$E\mathrm{tr}\{(\underline{S}^{-1}\tau(\underline{V}, \underline{S}) - \underline{\Sigma}^{-1})\underline{S}(\tau(\underline{V}, \underline{S})\underline{S}^{-1} - \underline{\Sigma}^{-1})Q_0\}$$

$$= E\mathrm{tr}\{[\underline{S}^{-1}\tau(\underline{V}, \underline{S})\underline{S}\,\tau(\underline{V}, \underline{S})\underline{S}^{-1} - 2\underline{\Sigma}^{-1}\underline{S}\,\tau(\underline{V}, \underline{S})\underline{S}^{-1} + \underline{\Sigma}^{-1}\underline{S}\,\underline{\Sigma}^{-1}]Q_0\}$$

$$= E\mathrm{tr}\{[\underline{S}^{-1}\tau(\underline{V}, \underline{S})\underline{S}\,\tau(\underline{V}, \underline{S})\underline{S}^{-1} - 2\underline{\Sigma}^{-1}\underline{S}\,\tau(\underline{V}, \underline{S})\underline{S}^{-1}]Q_0\} + K\mathrm{tr}(\underline{\Sigma}^{-1}Q_0) \tag{2.2.4}$$

To obtain (2.2.4), note that $\mathrm{tr}(\underline{\Sigma}^{-1}\underline{S}\,\tau(\underline{V}, \underline{S})\underline{S}^{-1}Q_0) = \mathrm{tr}(\underline{\Sigma}^{-1}\underline{S}\,\tau(\underline{V}, \underline{S})\underline{S}^{-1}Q_0)^T = \mathrm{tr}(Q_0\underline{S}^{-1}\tau(\underline{V}, \underline{S})\underline{S}\,\underline{\Sigma}^{-1}) = \mathrm{tr}(\underline{S}^{-1}\tau(\underline{V}, \underline{S})\underline{S}\,\underline{\Sigma}^{-1}Q_0)$, and use $E(\underline{S}) = K\underline{\Sigma}$.

From (2.2.2), (2.2.3) and (2.2.4), one has

$$EL_Q(\underline{\theta}, \widehat{\underline{\theta}}^{EB})$$

$$= \quad K\mathrm{tr}(\underline{V}\,\underline{Q})+E\mathrm{tr}\{[\underline{S}^{-1}\underline{\tau}(\underline{V},\underline{S})\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}-2\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}]\underline{Q}_0\} \qquad (2.2.5)$$

Note that $R_{\underline{Q}}(\underline{\theta},\underline{X}) = E_{\underline{\theta}}L_{\underline{Q}}(\underline{\theta},\underline{X}) = K\mathrm{tr}(\underline{V}\,\underline{Q})$ for all $\underline{\theta}$. Using the Wishart identity as given in Lemma 1.3.3,

$$E(\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1})$$

$$= \quad E(2\underline{D}^{\underline{S}}[\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}]+(K-p-1)\underline{S}^{-1}\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1})$$

$$= \quad E(2\underline{D}^{\underline{S}}[\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}]+(K-p-1)\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}) \qquad (2.2.6)$$

It follows from (2.2.5) and (2.2.6) that

$$EL_{\underline{Q}}(\underline{\theta},\widehat{\underline{\theta}}^{EB})$$

$$= \quad \tilde{E}R_{\underline{Q}}(\underline{\theta},\widehat{\underline{\theta}}^{EB})$$

$$= \quad \tilde{E}R_{\underline{Q}}(\underline{\theta},\underline{X})+\tilde{E}E_{\underline{\theta}}\mathrm{tr}\{[\underline{S}^{-1}\underline{\tau}(\underline{V},\underline{S})\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}-4\underline{D}^{\underline{S}}(\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}]$$

$$2(K-p-1)\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}]\underline{Q}_0\} \qquad (2.2.7)$$

The conditional distribution of $\underline{S} = \sum_{i=1}^{K}\underline{X}_i\underline{X}_i^T$ given $\underline{\theta}$ depends on $\underline{\theta}$ only through $\sum_{i=1}^{K}\underline{\theta}_i\underline{\theta}_i^T = \underline{\theta}\,\underline{\theta}^T = \underline{\Theta}$(say). Also, the family of distributions of $\underline{\Theta}$ being Wishart $W(p, K, \underline{\Delta})$ is complete for $\underline{\Delta}$ (or $\underline{\Sigma}$). Thus, (2.2.1) follows immediately from (2.2.7).

Remark 2.2.1 The technique of proof used in Theorem 2.2.1 is borrowed from Efron and Morris (1976b) (see their Theorem 2). They compute the risk of a special choice of $\widehat{\underline{\theta}}^{EB}$ when $\underline{V} = \underline{Q} = \underline{I}_p$. The more general (2.2.1) is available because of the powerful Wishart identity given in Lemma 1.3.3

Next, we will consider two classes of EB estimators of the form (2.1.8), first, $\widehat{\underline{\theta}}^{EB(1)} = (\underline{I}_p - r(\underline{V}, \underline{S})\underline{V}\,\underline{S}^{-1})\underline{X}$, so that $\underline{r}(\underline{V}, \underline{S}) = r(\underline{V}, \underline{S})\underline{I}_p$, where $r(\underline{V}, \underline{S})$ being a scalar in (2.1.8). Second, $\widehat{\underline{\theta}}^{EB(2)} = (\underline{I}_p - \underline{V}\,\underline{S}^{-1}\underline{\varphi}(\underline{S}))\underline{X}$, so that $\underline{r}(\underline{V}, \underline{S}) = \underline{\varphi}(\underline{S})$ in (2.1.8), where $\underline{\varphi}(\underline{S})$ is symmetric, a function of $\underline{S}$ only and depends on $\underline{X}$ only through $\underline{S}$. It is shown in Theorem 2.2.2 and Theorem 2.2.3 that $\widehat{\underline{\theta}}^{EB(1)}$ and $\widehat{\underline{\theta}}^{EB(2)}$ dominate $\underline{X}$ under the loss (2.1.5) and under certain conditions on $r(\underline{V}, \underline{S})$ and $\underline{\varphi}(\underline{S})$, respectively.

Corollary 2.2.1 Under the L_Q loss,

$$\widehat{\underline{\theta}}^{EB(1)} = (\underline{I}_p - r(\underline{V}, \underline{S})\underline{V}\,\underline{S}^{-1})\underline{X}, \tag{2.2.8}$$

has risk

$$R_Q(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)})$$

$$= R_Q(\underline{\theta}, \underline{X}) + E_{\underline{\theta}}\mathrm{tr}\{[r^2(\underline{V}, \underline{S}) - 2(K-p-1)r(\underline{V}, \underline{S})]\underline{S}^{-1}\underline{Q}_0 - 4[\underline{D}^{\underline{S}}\, r(\underline{V}, \underline{S})]\underline{Q}_0\} \tag{2.2.9}$$

Proof of Corollary 2.2.1 Let $\underline{r}(\underline{V}, \underline{S}) = r(\underline{V}, \underline{S})\underline{I}_p$ in Theorem 2.2.1. Hence (2.2.9) follows directly from (2.2.1).

In view of (2.2.9), an unbiased estimator of the risk difference $R_Q(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)}) - R_Q(\underline{\theta}, \underline{X})$ is given by

$$h_Q^{(1)}(\underline{V}, \underline{S})$$

$$= \mathrm{tr}\{[r^2(\underline{V}, \underline{S}) - 2(K-p-1)r(\underline{V}, \underline{S})]\underline{S}^{-1}\underline{Q}_0 - 4[\underline{D}^{\underline{S}}\, r(\underline{V}, \underline{S})]\underline{Q}_0\}$$

$$= \mathrm{tr}\{\underline{Q}_0^{1/2}[(r^2(\underline{V}, \underline{S}) - 2(K-p-1)r(\underline{V}, \underline{S}))\underline{S}^{-1} - 4\,(\underline{D}^{\underline{S}}\, r(\underline{V}, \underline{S}))]\underline{Q}_0^{1/2}\}, \tag{2.2.10}$$

where $\underline{Q}_0 = \underline{Q}_0^{1/2}\underline{Q}_0^{1/2}$ since $\underline{Q}_0^{1/2}$ is symmetric and p.d. It follows from Corollary 2.2.1 that to find EB estimator $\hat{\underline{\theta}}^{EB(1)}$ improving on $\underline{X}$, it suffices to find solutions for $h_{\underline{Q}}^{(1)}(\underline{V}, \underline{S}) \le 0$ with $P_{\underline{\theta}}(h_{\underline{Q}}^{(1)}(\underline{V}, \underline{S}) < 0) > 0$ for all $\underline{\theta}$. Using (2.2.10), it suffices to find solutions for the matrix inequality $\underline{H}^{(1)}(\underline{V}, \underline{S}) \le \underline{0}$, where

$$\underline{H}^{(1)}(\underline{V}, \underline{S}) = [r^2(\underline{V}, \underline{S}) - 2(K-p-1)r(\underline{V}, \underline{S})]\underline{S}^{-1} - 4[\underline{D}^{\underline{S}}\, r(\underline{V}, \underline{S})] \tag{2.2.11}$$

Recall that we write a matrix $\underline{T} \le \underline{0}$ if $\underline{T}$ is non-positive definite (n.p.d.) and $\underline{T} \ge \underline{0}$ if $\underline{T}$ is non-negative definite (n.n.d.). The following theorem provides a set of sufficient conditions under which $\hat{\underline{\theta}}^{EB(1)}$ will dominate $\underline{X}$.

Theorem 2.2.2 Let $\underline{S} = ((s_{ij}))$ and $r(\underline{V}, \underline{S})$ be a differentiable function with respect to s_{ij} for each $i, j = 1, \ldots, p$, then under the loss (2.1.5), $R_{\underline{Q}}(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}) < R_{\underline{Q}}(\underline{\theta}, \underline{X})$ if

(i) $0 < r(\underline{V}, \underline{S}) < 2(K-p-1)$,

(ii) $\underline{D}^{\underline{S}}\, r(\underline{V}, \underline{S})$ is n.n.d.

Proof of Theorem 2.2.2 Since $\underline{S}$ is p.d., assumptions (i) and (ii), and (2.2.11) imply that $\underline{H}^{(1)}(\underline{V}, \underline{S}) \le \underline{0}$. Hence (2.2.10) is negative. Therefore, the proof of Theorem 2.2.2 is complete.

Consider the special case $r(\underline{V}, \underline{S}) = r(F)$, where $F = tr(\underline{S}\,\underline{V}^{-1})$. The following corollary to Theorem 2.2.2 shows $\hat{\underline{\theta}}^{EB(1)}$ dominates $\underline{X}$ under the loss (2.1.5) and conditions similar to those proved in Baranchik (1970) for EB estimators of vector normal means.

Corollary 2.2.2 Let $r(\underline{V}, \underline{S}) = r(F)$, where $F = \text{tr}(\underline{S}\,\underline{V}^{-1})$, and $r(F)$ is differentiable. Then under the loss (2.1.5) $\hat{\underline{\theta}}^{EB(1)}$ dominates $\underline{X}$ if

(i) $0 < r(F) < 2(K-p-1)$

(ii) $r(F)$ is non-decreasing in F

Proof of Corollary 2.2.2 Let $\underline{V}^{-1} = ((v^{ij}))$. In view of condition (ii) in Theorem 2.2.2, $\underline{D}^S r(F) = r'(F)(\underline{D}^S F) = r'(F)((d_{ij}^S F))$ and $d_{ij}^S F = 1/2(1+\delta_{ij})\frac{\partial}{\partial s_{ij}}(\sum_{k=1}^{p}\sum_{m=1}^{p} s_{km} v^{mk}) = v^{ij}$, δ_{ij}'s being the Kronecker deltas. It follows from assumption (ii) that $\underline{D}^S r(F) = r'(F)\cdot\underline{V}^{-1} \geq \underline{0}$. The proof of the corollary is complete.

Remark 2.2.2 It is important to observe that, when $p = 1$, Corollary 2.2.2 reduces to the results in Baranchik (1970) for independent normal random variables which have a common known variance. His results on unknown covariance will be extended to the matrix case with unknown covariance matrix $\underline{V}$ in Chapter 3.

A matrix extension of the James Stein (1961) estimator is given in the next example.

Example 2.2.1 Suppose $r(\underline{V}, \underline{S}) = a$ is constant. Following (2.2.9), one gets

$$R_Q(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}) - R_Q(\underline{\theta}, \underline{X}) = E_{\underline{\theta}}\text{tr}\{a[a-2(K-p-1)]Q_0^{1/2}\underline{S}^{-1}Q_0^{1/2}\} \qquad (2.2.12)$$

Hence, the EB estimator $(I_p - a\underline{V}\,\underline{S}^{-1})\underline{X}$ dominates $\underline{X}$, if $0 < a < 2(K-p-1)$. The optimal choice of a is $K-p-1$, which leads to the matrix extension of James Stein estimator $(I_p - (K-p-1)\underline{V}\,\underline{S}^{-1})\underline{X}$.

The rest of this section concentrates on $\widehat{\underline{\theta}}^{EB(2)}$. We will propose several $\underline{g}(\underline{S})$ for $\widehat{\underline{\theta}}^{EB(2)}$ such that $\widehat{\underline{\theta}}^{EB(2)}$ is better than $\underline{X}$ under the L_Q loss. The risk $R_Q(\underline{\theta}, \widehat{\underline{\theta}}^{EB(2)})$ of $\widehat{\underline{\theta}}^{EB(2)}$ is provided below.

Corollary 2.2.3 Under the L_Q loss,

$$\widehat{\underline{\theta}}^{EB(2)} = (\underline{I}_p - \underline{V}\,\underline{S}^{-1}\underline{g}(\underline{S}))\underline{X}, \tag{2.2.13}$$

has risk

$$\begin{aligned} & R_Q(\underline{\theta}, \widehat{\underline{\theta}}^{EB(2)}) \\ = \; & R_Q(\underline{\theta}, \underline{X}) + E_{\underline{\theta}}\mathrm{tr}\{[\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1} - 4\underline{D}^{S}[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}] \\ & - 2(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}]\underline{Q}_0\} \end{aligned} \tag{2.2.14}$$

Proof of Corollary 2.2.3 Equation (2.2.14) follows directly from (2.2.1) by replacing $\underline{s}(\underline{V}, \underline{S})$ with $\underline{g}(\underline{S})$.

In view of (2.2.14) an unbiased estimator of the risk difference $R_Q(\underline{\theta}, \widehat{\underline{\theta}}^{EB(2)}) - R_Q(\underline{\theta}, \underline{X})$ is given by

$$\begin{aligned} & h_Q^{(2)}(\underline{V}, \underline{S}) \\ = \; & E_{\underline{\theta}}\mathrm{tr}\{[\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1} - 4\underline{D}^{S}[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}] - 2(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}]\underline{Q}_0\} \\ = \; & E_{\underline{\theta}}\mathrm{tr}\{\underline{Q}_0^{1/2}[\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1} - 4\underline{D}^{S}[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}] - 2(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}]\underline{Q}_0^{1/2}\} \end{aligned} \tag{2.2.15}$$

Note that $\hat{\underline{\theta}}^{EB(2)}$ dominates $\underline{X}$ if $h_Q^{(2)}(\underline{V}, \underline{S})) \leq 0$ with $P_{\underline{\theta}}(h_Q^{(2)}(\underline{V}, \underline{S}) < 0) > 0$ for all $\underline{\theta}$. From (2.2.15), it suffices to show the matrix $\underline{H}^{(2)}(\underline{S}) \leq \underline{0}$, where

$$\underline{H}^{(2)}(\underline{S}) = \underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1} - 4\underline{D}^{S}[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}] - 2(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1} \tag{2.2.16}$$

Consider the particular choice $\underline{g}(\underline{S}) = \underline{P}\,\underline{\Psi}(\underline{\lambda})\underline{P}^T$, where $\underline{\lambda} = (\lambda_1, \ldots, \lambda_p)^T$ denotes the vector of eigen-values of $\underline{S}$, and $\underline{\Psi}(\underline{\lambda})$ is a diagonal matrix with ith diagonal element equal to $\Psi_i(\underline{\lambda}) > 0$, $i = 1, \ldots, p$. Also, let $\underline{P}$ be an orthogonal matrix with its column vectors equal to the orthonormal eigen vectors of $\underline{S}$. Since $\underline{S} = \underline{P}\,\underline{\Delta}\,\underline{P}^T$, where $\underline{\Delta}$ is a diagonal matrix with its diagonal elements equal to the eigen-values of $\underline{S}$, it follows from (2.2.16) that with the given choice of $\underline{g}(\underline{S})$

$$\underline{H}^{(2)}(\underline{S}) = \underline{P}\,\underline{\Psi}^2(\underline{\lambda})\underline{\Delta}^{-1}\underline{P}^T - 4\underline{D}^{S}[\underline{P}\,\underline{\Psi}(\underline{\lambda})\underline{P}^T] - 2(K-p-1)\underline{P}\,\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}\underline{P}^T \tag{2.2.17}$$

From Haff (1984), one gets

$$\underline{D}^{S}[\underline{P}\,\underline{\Psi}(\underline{\lambda})\underline{P}^T] = \underline{P}\,\underline{\Psi}^{(1)}(\underline{\lambda})\underline{P}^T, \tag{2.2.18}$$

where $\underline{\Psi}^{(1)}(\underline{\lambda})$ is a diagonal matrix with ith diagonal element equal to

$$\Psi_i^{(1)}(\underline{\lambda}) = \frac{1}{2}\sum_{j(\neq i)}[\Psi_i(\underline{\lambda}) - \Psi_j(\underline{\lambda})]/(\lambda_i - \lambda_j) + \partial\Psi_i(\underline{\lambda})/\partial\lambda_i \tag{2.2.19}$$

From (2.2.16) and (2.2.17), $\underline{H}^{(2)}(\underline{S}) \leq \underline{0}$ is equivalent to $\underline{u}(\underline{\lambda}) \leq \underline{0}$, where $\underline{u}(\underline{\lambda})$ is a diagonal matrix with ith diagonal element equal to

$$u_i(\underline{\lambda}) = \Psi_i^2(\underline{\lambda})\lambda_i^{-1} - 4\Psi_i^{(1)}(\underline{\lambda}) - 2(K-p-1)\Psi_i(\underline{\lambda})\lambda_i^{-1} \tag{2.2.20}$$

In the remainder of the section, we find solutions of $u_i(\underline{\lambda}) < 0$ for each $i = 1, \ldots, p$, obtaining thereby a class of EB estimators dominating $\underline{X}$.

The following theorem shows that Zheng's (1986) Theorem 2 continues to hold under the general loss $L_{\underline{Q}}$ and variance matrix $\underline{V}$. Zheng (1986) proved the result for $\underline{Q} = \underline{V} = \underline{I}_p$.

Theorem 2.2.3 Suppose that

(i) $0 < \Psi_i(\underline{\lambda}) < 2(K-p-1)$ for each $i = 1, \ldots, p$,

(ii) $\partial\Psi_i(\underline{\lambda})/\partial\lambda_i \geq 0$ for each $i = 1, \ldots, p$,

(iii) $\Psi_i(\underline{\lambda})$'s are similarly ordered with the λ_i's, that is,

$(\Psi_i(\underline{\lambda})-\Psi_j(\underline{\lambda}))(\lambda_i-\lambda_j) \geq 0$ for every $1 \leq i, j \leq p$.

Then $u_i(\underline{\lambda}) < 0$ for each $i = 1, \ldots, p$.

Proof of Theorem 2.2.3 In view of assumptions (ii) and (iii), $\Psi_i^{(1)}(\underline{\lambda}) \geq 0$ for each $i = 1, \ldots, p$. Hence, from (2.2.20)

$$u_i(\underline{\lambda}) \leq \Psi_i(\underline{\lambda})[\Psi_i(\underline{\lambda})-2(K-p-1)]\lambda_i^{-1} < 0, \tag{2.2.21}$$

using assumption (i). The proof of Theorem 2.2.3 is complete.

We provide two examples illustrating the use of Theorem 2.2.3

Example 2.2.2 Stein (1974), and Efron and Morris (1976b) have shown that if $\underline{S}^{-1}\underline{\varphi}(\underline{S}) = a\underline{S}^{-1}+(b/\mathrm{tr}(\underline{S}))\underline{I}_p$, for $\underline{V} = \underline{Q} = \underline{I}_p$, then $(\underline{I}_p-\underline{S}^{-1}\underline{\varphi}(\underline{S}))\underline{X}$ dominates $\underline{X}$ for certain choices of the constants a and b. Note that one can write $\underline{S}^{-1}\underline{\varphi}(\underline{S}) = \underline{P}\underline{\Delta}^{-1}\underline{\Psi}(\underline{\lambda})\underline{P}^T$, where $\underline{\Delta}^{-1}\underline{\Psi}(\underline{\lambda})$ is a diagonal matrix with ith diagonal element equal to

$$\lambda_i^{-1}\Psi_i(\underline{\lambda}) = \lambda_i^{-1}(a+b\lambda_i/\sum_{j=1}^{p}\lambda_j) \tag{2.2.22}$$

Then if $a \geq 0$, $b \geq 0$ and $0 < a+b < 2(K-p-1)$, all the three conditions of Theorem 2.2.3 are met with probability 1. Consequently for such a pair (a, b), the estimator $(\underline{I}_p-\underline{V}[a\underline{S}^{-1}+(b/\mathrm{tr}(\underline{S}))\underline{I}_p])\underline{X}$ dominates $\underline{X}$ for every L_Q loss given in (2.1.5). In particular, if $a = K-p-1$ then for every $b \in [0, K-p-1)$, the dominance over $\underline{X}$ holds. Also, if $b = 0$, then $(\underline{I}_p-a\underline{V}\,\underline{S}^{-1})\,\underline{X}$ dominates $\underline{X}$ for every $a \in (0, 2(K-p-1))$.

Example 2.2.3 Suppose now $\underline{S}^{-1}\underline{\varphi}(\underline{S}) = (K-p-1)\underline{S}^{-1}+[b/\mathrm{tr}(\underline{S}^2)]\underline{S}$ with $b \in [0, K-p-1)$. Then,

$$\Psi_i(\underline{\lambda}) = (K-p-1)+b\,\lambda_i^2/\sum_{j=1}^{p}\lambda_j^2 \tag{2.2.23}$$

It is easy to verify conditions (i)-(iii) of Theorem 2.2.3 for this particular choice of $\underline{\Psi}(\underline{\lambda})$, and accordingly, $(\underline{I}_p-\underline{V}\,\underline{S}^{-1}\underline{\varphi}(\underline{S}))\underline{X}$ dominates $\underline{X}$ for every L_Q loss.

It is possible, however, to obtain estimators of the form (2.2.13) which dominate $\underline{X}$ even though the conditions of Theorem 2.2.3 are not met. Some examples follow.

Example 2.2.4 Consider the case when b, c and d are constants,

$$\underline{\Psi}(\underline{\lambda}) = b(1+cU_1^d)^{-1}\underline{1}_p \tag{2.2.24}$$

and

$$U_1 = \prod_{j=1}^{p}\lambda_j^{1/p} \tag{2.2.25}$$

is the geometric mean of the eigen-values of $\underline{S}$. Then, condition (ii) of Theorem 2.2.3 is violated, but

$$u_i(\underline{\lambda}) = \lambda_1^{-1}(1+cU_1^d)^{-2}b[b-2(K-p-1)-2cU_1^d(K-p-1-2dp^{-1})] < 0 \qquad (2.2.26)$$

if (a) $0 < b < 2(K-p-1)$, (b) $c > 0$ and (c) $0 \leq d \leq p(K-p-1)/2$. Thus, $(\underline{I}_p-b(1+cU_1^d)^{-1}\underline{V}\,\underline{S}^{-1})\underline{X}$ dominates $\underline{X}$ for every L_Q loss given in (2.1.5) if (a)-(c) hold.

Example 2.2.5 Suppose this time

$$\underline{\Psi}(\underline{\lambda}) = b(1+cU_2^d)^{-1}\underline{I}_p \qquad (2.2.27)$$

and

$$U_2 = \sum_{j=1}^{p}\lambda_j/p \qquad (2.2.28)$$

is the arithmetic mean of the eigen values of $\underline{S}$. Once again, condition (ii) of Theorem 2.2.3 is violated but

$$\begin{aligned} u_i(\underline{\lambda}) &= \lambda_1^{-1}(1+cU_2^d)^{-2}b[b-2(K-p-1)-2cU_2^d(K-p-1-2d\lambda_i(pU_2)^{-1})] \\ &\leq \lambda_1^{-1}(1+cU_2^d)^{-2}b[b-2(K-p-1)-2cU_2^d(K-p-1-2d)] < 0 \qquad (2.2.29) \end{aligned}$$

if (a) $0 < b < 2(K-p-1)$, (b) $c > 0$, (c) $0 \leq d \leq (K-p-1)/2$. Thus, $(\underline{I}_p-b(1+cU_2^d)^{-1}\underline{V}\,\underline{S}^{-1})\underline{X}$ dominates $\underline{X}$ for every L_Q loss given in (2.1.5) if (a)-(c) hold. One can also use

$$\underline{\Psi}(\underline{\lambda}) = b(1+cU_3^d)^{-1}\underline{I}_p, \qquad (2.2.30)$$

where

$$U_3 = p(\sum_{j=1}^{p}\lambda_j^{-1})^{-1}, \qquad (2.2.31)$$

the harmonic mean of the eigen-values of $\underline{S}$. Then, $(\underline{I}_p-b(1+cU_3^{d})^{-1}\underline{V}\,\underline{S}^{-1})\underline{X}$ dominates $\underline{X}$ for every L_Q loss under conditions (a)-(c) of the example

Example 2.2.6 We consider now a class of estimators $\{\underline{S}^{-1}\underline{g}(\underline{S})\}_{\alpha}$ of $\underline{\Sigma}^{-1}$, where

$$\{\underline{S}^{-1}\underline{g}(\underline{S})\}_{\alpha} = (1-\alpha)\{(K-p-1)\underline{S}^{-1}+(b/\mathrm{tr}(\underline{S}))\underline{I}_p\}+\alpha\{((Kp-2)/\mathrm{tr}(\underline{S}))\underline{I}_p\}, \tag{2.2.32}$$

$0 \le \alpha \le 1$ An alternate representation of $\{\underline{S}^{-1}\underline{g}(\underline{S})\}_{\alpha} = \underline{P}\,\underline{\Delta}^{-1}\underline{\Psi}(\underline{\lambda})\,\underline{P}^T$, where $\underline{\Psi}(\underline{\lambda})$ is a diagonal matrix with ith diagonal element equal to

$$\Psi_i(\underline{\lambda}) = a+b_0\lambda_i/\sum_{j=1}^{p}\lambda_j, \tag{2.2.33}$$

where $a = (1-\alpha)(K-p-1)$, $b_0 = (1-\alpha)b+\alpha(Kp-2)$

Note that from Example 2.2.2, $(\underline{I}_p-\underline{V}\{\underline{S}^{-1}\underline{g}(\underline{S})\}_{\alpha})\underline{X}$ dominates $\underline{X}$ for all L_Q if $0 < a+b_0 < 2(K-p-1)$ i.e., $0 < b_0 < (1+\alpha)(K-p-1)$ or $0 < b < [(1+\alpha)(K-p-1)-\alpha(Kp-2)]/(1-\alpha)$ Note also that the above holds only if $\alpha \in (0, (K-p-1)/[(p-1)(K+1)])$

Remark 2.2.3 Efron and Morris (1976b) chose $b = p^2+p-2$ and showed the dominance of the corresponding $(\underline{I}_p-\{\underline{S}^{-1}\underline{g}(\underline{S})\}_{\alpha})\underline{X}$ over $\underline{X}$ under the squared error loss and $\underline{V} = \underline{I}_p$ We do not have such a strong result regarding dominance of $(\underline{I}_p-\underline{V}\{\underline{S}^{-1}\underline{g}(\underline{S})\}_{\alpha})\underline{X}$ over $\underline{X}$ simultaneously for all L_Q. However we shall see from Section 2.3 that a very appropriate choice of b is p-1 With this particular choice of b, it follows from (2.2.20) that

$$\begin{aligned} u_i(\underline{\lambda}) &= (a^2-2(K-p-1)a)\lambda_i^{-1}+(2ab_0-2(K-p-1)b_0-2(p-1)b_0-4b_0)/(\sum_{j=1}^{p}\lambda_j) \\ &\quad +b_0^2\lambda_i/(\sum_{j=1}^{p}\lambda_j)^2+4b_0\lambda_i/(\sum_{j=1}^{p}\lambda_j)^2 \\ &\le [a^2-2(K-p-1)a-2(K-2-a)b_0+b_0^2]/(\sum_{j=1}^{p}\lambda_j) \end{aligned} \tag{2.2.34}$$

Note that the right hand side of (2.2.34) is negative for $\alpha \in [0, (K-p-1)(K^2(p-1))^{-1}\{1+[1+K^2(K-p-1)^{-2}((K-p-1)^2+(p-1)^2)]^{1/2}\}]$, and the function is minimized at $\alpha_0 = (K-p-1)/[K^2(p-1)]$. We propose, therefore, the estimator

$$\hat{\underline{\theta}}_{\alpha_0}(\underline{X})$$

$$= (\underline{I}_p-\underline{V}[(1-\alpha_0)\{(K-p-1)\underline{S}^{-1}+((p-1)/tr(\underline{S}))\underline{I}_p\}+\alpha_0\{(Kp-2)/tr(\underline{S})\}\underline{I}_p])\underline{X}. \qquad (2.2.35)$$

2.3 Improvement On $(\underline{I}_p-r(\underline{V},\underline{S})\underline{V}\,\underline{S}^{-1})\underline{X}$

Consider $\hat{\underline{\theta}}^{EB(1)}(\underline{X}) = (\underline{I}_p-r(\underline{V},\underline{S})\underline{V}\,\underline{S}^{-1})\underline{X}$, where $r(\underline{V},\underline{S})$ is differentiable with respect to s_{ij} for each $i, j = 1, \ldots, p$, and let $\hat{\underline{\theta}}^{EB(3)}(\underline{X}) = (\underline{I}_p-r(\underline{V},\underline{S})\underline{V}\,\underline{S}^{-1}-\underline{V}\,\underline{S}^{-1}\underline{f}(\underline{S}))\underline{X}$, where $\underline{f}(\underline{S}) = \underline{P}\,\underline{\Psi}(\underline{\lambda})\,\underline{P}^T$. $\underline{\Psi}(\underline{\lambda})$ is a diagonal matrix with ith diagonal element equal to $\Psi_i(\underline{\lambda}) > 0$, $i = 1, \ldots, p$, as before $\underline{P}$ is an orthogonal matrix with ith column vectors equal to the orthonormal eigen-vectors of $\underline{S}$, and $\underline{\lambda} = (\lambda_1, \ldots, \lambda_p)^T$ denotes the vector of eigen-values of $\underline{S}$. Then from (2.2.1), one gets

$$R_Q(\underline{\theta}, \hat{\underline{\theta}}^{EB(3)}(\underline{X}))-R_Q(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}(\underline{X}))$$

$$= E_{\underline{\theta}} tr\{[\underline{S}^{-1}\underline{f}(\underline{S})\underline{S}\,\underline{f}(\underline{S})\underline{S}^{-1}+r(\underline{V},\underline{S})(\underline{S}^{-1}\underline{f}(\underline{S})+\underline{f}(\underline{S})\underline{S}^{-1})-4\underline{D}^{S}(\underline{S}\,\underline{f}(\underline{S})\underline{S}^{-1})$$

$$-2(K-p-1)\underline{f}(\underline{S})\underline{S}^{-1}]Q_0\}$$

$$= E_{\underline{\theta}} tr(Q_0^{1/2}\underline{P}[\underline{\Psi}^2(\underline{\lambda})\Delta^{-1}-2(K-p-1-r(\underline{V},\underline{S}))\underline{\Psi}(\underline{\lambda})\Delta^{-1}-4\underline{\Psi}^{(1)}(\underline{\lambda})]\underline{P}^T Q_0^{1/2}) < 0$$

for all $\underline{\theta}$ if $\Delta_{Q_i}(\underline{\lambda}) < 0$ for each $i = 1, \ldots, p$, where

$$\Delta_{Q_i}(\underline{\lambda}) = \Psi_i^2(\underline{\lambda})\lambda_i^{-1}-2(K-p-1-r(\underline{V},\underline{S}))\Psi_i(\underline{\lambda})\lambda_i^{-1}-2\sum_{j(\neq i)}\{\Psi_i(\underline{\lambda})-\Psi_j(\underline{\lambda})\}/(\lambda_i-\lambda_j)$$

$$-4[\partial\Psi_i(\underline{\lambda})/\partial\lambda_i] \qquad (2.3.1)$$

We now prove a general theorem providing solutions to $\Delta_{Qi}(\underline{\lambda}) < 0$ for each $i = 1, \ldots, p$.

Theorem 2.3.1 Let $0 < r(\underline{V}, \underline{S}) < (K-2)$ and $r(\underline{V}, \underline{S})$ is differentiable with respect to s_{ij} for each $i, j = 1, \ldots, p$. Suppose $\Psi_i(\underline{\lambda}) = \lambda_i t(U)/(\sum_{j=1}^{p}\lambda_j)$, $i = 1, \ldots, p$, where U is a function of the λ_i's. Suppose that

(i) $U(\underline{\lambda})$ is non-decreasing in λ_i for each $i = 1, \ldots, p$,

(ii) $t(U)$ is non-decreasing in U,

(iii) $0 < t(U) < 2(K-r(\underline{V}, \underline{S})-2)$

Then $\Delta_{Qi}(\underline{\lambda}) < 0$ for each $i = 1, \ldots, p$. Hence, $\hat{\underline{\theta}}^{EB(3)}(\underline{X})$ dominates $\hat{\underline{\theta}}^{EB(1)}(\underline{X})$ under the L_Q loss.

Proof of Theorem 2.3.1 It follows from (2.3.1) that

$$\begin{aligned}
\Delta_{Qi}(\underline{\lambda}) = {} & \lambda_i t^2(U)/(\sum_{j=1}^{p}\lambda_j)^2-2(K-p-1-r(\underline{V}, \underline{S}))t(U)/(\sum_{j=1}^{p}\lambda_j)-2(p-1)t(U)/(\sum_{j=1}^{p}\lambda_j) \\
& -4[t(U)/(\sum_{j=1}^{p}\lambda_j)-\lambda_i t(U)/(\sum_{j=1}^{p}\lambda_j)^2+(\lambda_i/\sum_{j=1}^{p}\lambda_j)t'(U)(\partial U/\partial\lambda_i)] \\
= {} & [t^2(U)+4t(U)]\lambda_i/(\sum_{j=1}^{p}\lambda_j)^2-2(K-r(\underline{V}, \underline{S}))t(U)/(\sum_{j=1}^{p}\lambda_j) \\
& -4(\lambda_i/\sum_{j=1}^{p}\lambda_j)t'(U)(\partial U/\partial\lambda_i) \\
\leq {} & [t^2(U)+4t(U)]\lambda_i/(\sum_{j=1}^{p}\lambda_j)^2-2(K-r(\underline{V}, \underline{S}))t(U)/(\sum_{j=1}^{p}\lambda_j) \\
\leq {} & [t^2(U)-2(K-r(\underline{V}, \underline{S})-2)t(U)]/(\sum_{j=1}^{p}\lambda_j) < 0 \qquad (2.3.2)
\end{aligned}$$

The first inequality in (2.3.2) uses parts (i) and (ii) of the assumption, while the last inequality uses part (iii). The proof of the theorem is complete.

Example 2.3.1 It follows as a special case of Theorem 2.3.1 that when $\tau(\underline{V}, \underline{S}) = a$ is constant, $(\underline{I}_p - a\underline{V}\,\underline{S}^{-1} - [t(U)/tr(\underline{S})]\underline{V})\underline{X}$ dominates $(\underline{I}_p - a\underline{V}\,\underline{S}^{-1})\underline{X}$ whenever $0 < a < (K-2)$ and $0 < t(U) < 2(K-a-2)$. For the particular choice of $a = K-p-1$, $0 < t(U) < 2(p-1)$. In Dey, Ghosh and Srinivasan (1985), the dominance holds for all $0 < t(U) < 2(p^2+p+2)$, but only under the squared error loss and $\underline{V} = \underline{I}_p$.

Example 2.3.2 If $t(U) = b$ is constant, and $a = K-p-1$, the first inequality of (2.3.2) becomes an equality. From the penultimate step of (2.3.2), it follows that the optimal choice of b is $p-1$.

Remark 2.3.1 It may be noted that the choice of

$$U = |\underline{S}|^{1/p} = \prod_{j=1}^{p}\lambda_j^{1/p},\ tr(\underline{S}) = \sum_{j=1}^{p}\lambda_j/p \text{ or } p/(tr(\underline{S}^{-1})) = p(\sum_{j=1}^{p}\lambda_j^{-1})^{-1} \qquad (2.3.3)$$

(the geometric, arithmetic and harmonic mean of the eigen values of $\underline{S}$) all satisfy condition (i) of Theorem 2.3.1.

2.4 Simulation Results

In this section, we compare the simulated risks of five different estimators of $\underline{\theta}$. These estimators are $\underline{X}$ and $\hat{\underline{\theta}}_{(i)} = (\underline{I}_p - \underline{V}\,\underline{S}^{-1}\tau_{(i)}(\underline{V}, \underline{S}))\underline{X}$ $(i = 1, 2, 3, 4)$ where

$$\tau_{(1)}(\underline{V}, \underline{S}) = (K-p-1)\underline{I}_p + \{(p-1)/tr(\underline{S})\}\underline{S}; \qquad (2.4.1)$$

$$\tau_{(2)}(\underline{V}, \underline{S}) = (K-p-1)\underline{I}_p + \{(p-1)/tr(\underline{S}^2)\}\underline{S}^2; \qquad (2.4.2)$$

$$\Sigma_{(3)}(\underline{V}, \underline{S}) = (K-p-1)\underline{I}_p \tag{2.4.3}$$

$$\Sigma_{(4)}(\underline{V}, \underline{S}) = (1-\alpha_0)\{(K-p-1)\underline{I}_p+[(p-1)/\mathrm{tr}(\underline{S})]\underline{S}\}+\alpha_0\{[(Kp-2)/\mathrm{tr}(\underline{S})]\underline{S}\} \tag{2.4.4}$$

where $\alpha_0 = (K-p-1)/[K^2(p-1)]$

Throught, we have considered the loss

$$L_{\underline{Q}}(\underline{\theta}, \underline{a}) = \sum_{i=1}^{K} (\underline{\theta}_i-\underline{a}_i)^T \underline{Q}(\underline{\theta}_i-\underline{a}_i)/\mathrm{tr}(\underline{Q}\,\underline{V}) \tag{2.4.5}$$

The (K, p) combinations chosen are (4, 2), (8, 2), (12, 2), (5, 3), (9, 3) and (13, 3). For each p, 1000 samples are generated from $N_p(\underline{\theta}_i, \underline{V})$.

For p = 2, we have taken $\underline{\theta}_i = \underline{0}$ or $\underline{\theta}_i = (10, 35)^T$ for each i = 1, …, K. For each p = 3, our choice is $\underline{\theta}_i = \underline{0}$ or $\underline{\theta}_i = (6, 10, 50)^T$ for each i = 1, …, K. Corresponding to a given situation, for each (K, 2) pair, there are four ($\underline{Q}$, $\underline{V}$) combinations with each $\underline{Q}$ or $\underline{V}$ equal to $\underline{I}_2$ or $\underline{M}_2 = \mathrm{diag}(0.01, 1.00)$. For each (K, 3) pair, there are once again four ($\underline{Q}$, $\underline{V}$) combinations with each $\underline{Q}$ or $\underline{V}$ equal to $\underline{I}_3$ or $\underline{M}_3 = \mathrm{diag}(0.01, 1.00, 100.00)$. Thus, in this situation, for each (K, p, $\underline{Q}$, $\underline{V}$, $\underline{\theta}_i$), the simulated risk is the average of 1000 expected losses. Tables 2.1 and 2.2 report the simulated risks for $\underline{\theta}_i = \underline{0}$ (p×1) while Tables 2.3 and 2.4 report the same when $\underline{\theta}_i = (10, 35)^T$ and $\underline{\theta}_i = (6, 10, 50)^T$, respectively.

An examination of Tables 2.1-2.4 reveals that like the usual James-Stein estimator, shrinkage estimators of matrix normal means perform the best when the true $\underline{\theta}_i$ equals the point towards which we want to shrink (in this case $\underline{0}$ (p×K)). As the true $\underline{\theta}_i$ gets further and further away from the shrinkage point, the risk improvement also becomes less and less. However, in the present study, the overall risk improvement seems to be quite impressive even when the $\underline{\theta}_i$'s are far away from zeros except when $\underline{V} = \underline{I}_2$ and $\underline{Q} = \underline{M}_2$ or $\underline{V} = \underline{I}_3$ and $\underline{Q} = \underline{M}_3$.

Another point to note is that for fixed p, $\underline{Q}$, and $\underline{V}$, the risk improvement of each estimator increases over K. The improvement from K = p+2 to K = p+6 is much more pronounced than the one from K = p+6 to K = p+10. Also, it may be noted that $\hat{\underline{\theta}}_{(1)}$, $\hat{\underline{\theta}}_{(2)}$ and $\hat{\underline{\theta}}_{(4)}$ all outperform $\hat{\underline{\theta}}_{(3)}$ in terms of risk improvement when the true $\underline{\theta}_i$ values are zeros, although for $\underline{\theta}_i$'s far away from zeros, all the four estimators are nearly at par. Finally, we observe that $\hat{\underline{\theta}}_{(4)}$ has a consistent edge over the remaining estimators when $\underline{\theta}_i$'s are zeros, while $\hat{\underline{\theta}}_{(4)}$ performs as well as the others when $\underline{\theta}_i$'s are far away from zeros. Thus, we recommend the use of $\hat{\underline{\theta}}_{(4)}$ in the present context.

TABLE 2.1

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_1 = \underline{0}$, $\underline{V}$ known)

	K	$\underline{X}$	RISKS OF $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	PERCENTAGE RISK IMPROVEMENTS BY $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_2$	4	4.0416	2.4212	2.3438	3.0128	2.2505	40.09	42.01	25.46	44.32
	8	8.1000	2.8399	2.7222	3.0984	2.5692	64.94	66.39	61.75	68.28
$\underline{Q} = \underline{I}_2$	12	11.9069	2.8216	2.7224	2.9743	2.5985	76.30	77.14	75.02	78.18
$\underline{V} = \underline{I}_2$	4	3.9507	2.3453	2.2674	2.9161	2.1809	40.64	42.61	26.19	44.80
	8	7.9782	2.7454	2.6304	2.9938	2.4816	65.59	67.03	62.48	68.89
$\underline{Q} = \underline{M}_2$	12	12.0327	2.8895	2.7904	3.0511	2.6610	75.99	76.81	74.64	77.89
$\underline{V} = \underline{M}_2$	4	3.9507	2.3767	2.3924	2.9161	2.3223	39.84	39.44	26.19	41.22
	8	7.9782	2.8341	2.8402	2.9938	2.7706	64.48	64.40	62.48	65.27
$\underline{Q} = \underline{I}_2$	12	12.0327	2.9628	2.9673	3.0511	2.9178	75.38	75.34	74.64	75.75
$\underline{V} = \underline{M}_2$	4	3.9489	2.3697	2.3854	2.9141	2.3148	39.99	39.59	26.20	41.38
	8	7.9758	2.8305	2.8366	2.9917	2.7662	64.51	64.43	62.49	65.32
$\underline{Q} = \underline{M}_2$	12	12.0352	2.9635	2.9680	3.0527	2.9180	75.38	75.34	74.64	75.75

TABLE 2.2

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_i = \underline{0}$, $\underline{V}$ known)

			RISKS OF				PERCENTAGE RISK IMPROVEMENTS BY			
	K	$\underline{X}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_3$	5	4.9894	2.9180	2.8248	3.8424	2.8282	41.52	43.38	22.99	43.32
	9	9.0453	3.5160	3.3312	3.9955	3.3458	61.13	63.17	55.83	63.01
$\underline{Q} = \underline{I}_3$	13	12.8496	3.5364	3.3598	3.8486	3.3839	72.48	73.86	70.05	73.67
$\underline{V} = \underline{I}_3$	5	5.0346	3.0273	2.9247	3.9591	2.9339	39.87	41.91	21.36	41.72
	9	9.0037	3.5135	3.3341	3.9845	3.3441	60.98	62.97	55.75	62.86
$\underline{Q} = \underline{M}_3$	13	13.0482	3.4953	3.3185	3.8272	3.3456	73.21	74.57	70.67	74.36
$\underline{V} = \underline{M}_3$	5	5.0346	2.6044	2.6395	3.9591	2.5825	48.27	47.57	22.36	48.71
	9	9.0037	3.3908	3.3988	3.9845	3.3322	62.34	62.25	55.75	62.99
$\underline{Q} = \underline{I}_3$	13	13.0482	3.4032	3.4064	3.8272	3.3365	73.92	73.89	70.67	74.43
$\underline{V} = \underline{M}_3$	5	5.0349	2.5927	2.6277	3.9604	2.5705	48.51	47.81	21.34	48.95
	9	9.0037	3.3849	3.3929	3.9844	3.3258	62.41	62.32	55.75	63.06
$\underline{Q} = \underline{M}_3$	13	13.0511	3.3993	3.4024	3.8274	3.3319	73.95	73.93	70.67	74.47

TABLE 2.3

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($P = 2$, $\underline{\theta}_1 = (10, 35)^T$, $\underline{V}$ known)

	K	RISKS OF $\underline{X}$					PERCENTAGE RISK IMPROVEMENTS BY			
			$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_2$	4	3.9310	3.3659	3.3663	3.3668	3.3753	14.38	14.37	14.35	14.14
	8	7.9867	5.5328	5.5329	5.5332	5.5403	30.73	30.72	30.72	30.63
$\underline{Q} = \underline{I}_2$	12	11.9198	7.4663	7.4664	7.4663	7.4779	37.36	37.36	37.36	37.26
$\underline{V} = \underline{I}_2$	4	3.8620	3.7683	3.7684	3.7690	3.7696	2.43	2.42	2.41	2.39
	8	8.1606	7.7177	7.7178	7.7183	7.7258	5.43	5.43	5.42	5.39
$\underline{Q} = \underline{M}_2$	12	12.0685	11.3034	11.3035	11.3031	11.3067	6.34	6.34	6.34	6.31
$\underline{V} = \underline{M}_2$	4	3.8620	2.8708	2.8708	2.8714	2.8876	25.67	25.66	25.65	25.23
	8	8.1606	3.7111	3.7111	3.7116	3.7389	54.52	54.52	54.52	54.18
$\underline{Q} = \underline{I}_2$	12	12.0685	3.9567	3.9567	3.9563	4.0049	67.21	67.21	67.22	66.81
$\underline{V} = \underline{M}_2$	4	3.8606	2.8609	2.8609	2.8616	2.8779	25.90	25.89	25.88	25.45
	8	8.1641	3.6754	3.6754	3.6759	3.7035	54.98	54.98	54.97	54.64
$\underline{Q} = \underline{M}_2$	12	12.0714	3.8887	3.8887	3.8883	3.9374	67.69	67.79	67.79	67.38

TABLE 2.4

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_i = (6, 10, 50)^T$, $\underline{V}$ known)

	K	$\underline{X}$	RISKS OF $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	PERCENTAGE RISK IMPROVEMENTS BY $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = I_3$	5	5.0817	4.3674	4.3681	4.3685	4.3694	14.06	14.04	14.03	14.02
	9	8.9558	5.7199	5.7202	5.7204	5.7169	36.13	36.13	36.13	36.17
$Q = I_3$	13	13.0744	6.9719	6.9722	6.9722	6.9814	46.67	46.67	46.67	46.60
$\underline{V} = I_3$	5	5.0664	5.0010	5.0010	5.0013	5.0012	1.29	1.29	1.29	1.29
	9	9.0422	8.7239	8.7240	8.7243	8.7242	3.54	3.54	3.54	3.54
$Q = M_3$	13	13.0520	12.5177	12.5177	12.5177	12.5181	4.09	4.09	4.09	4.09
$\underline{V} = M_3$	5	5.0664	3.8372	3.8419	3.8743	3.8440	24.26	24.17	23.53	24.13
	9	9.0442	3.8906	3.8932	3.9136	3.9011	56.98	56.95	56.73	56.87
$Q = I_3$	13	13.0520	4.1711	4.1727	4.1834	4.1713	68.04	68.03	67.95	68.04
$\underline{V} = M_3$	5	5.0654	3.8356	3.8403	3.8731	3.8421	24.28	24.18	23.54	24.14
	9	9.0457	3.8878	3.8903	3.9109	3.8984	57.02	56.99	56.76	56.90
$Q = M_3$	13	13.0520	4.1703	4.1720	4.1828	4.1604	68.05	68.04	67.95	68.05

CHAPTER 3
EMPIRICAL BAYES MINIMAX ESTIMATORS OF MATRIX NORMAL MEANS FOR ARBITRARY QUADRATIC LOSS AND UNKNOWN COVARIANCE MATRIX IN THE ONE SAMPLE PROBLEM

3.1 Introduction

The investigation of the previous chapter has mainly concentrated on the situation when the covariance $\underline{V}$ of $\underline{X}_i$'s is known. In a more realistic situation, $\underline{V}$ will not be unknown, but we have supplementary data to estimate it. This is the situation in most multivariate normal models. Suppose we continue to consider the same set up as given in (2.1.1) and (2.1.3) but this time $\underline{V}$ is unknown and an estimator $\widehat{\underline{V}}$ a symmetric matrix independent of $\underline{X}$, is available as an estimate of $\underline{V}$. It is desired to estimate $\underline{\theta}$ under the loss

$$L_{Q^*}(\underline{\theta}, \underline{a}) = \sum_{i=1}^{K} (\underline{\theta}_i - \underline{a}_i)^T Q^* (\underline{\theta}_i - \underline{a}_i), \tag{3.1.1}$$

where $Q^* = \underline{V}^{-1/2} Q \, \underline{V}^{-1/2}$, $\underline{V} = \underline{V}^{1/2}\underline{V}^{1/2}$, and Q is p.d. and known. It can be shown that the usual estimator $\underline{X}$ of $\underline{\theta}$ has constant risk $K\mathrm{tr}(Q)$ for all $\underline{\theta}$ under the loss (3.1.1) instead of the value $K\mathrm{tr}(\underline{V}Q)$ under the loss (2.1.5). In this way $\underline{X}$ is a constant risk minimax estimator of $\underline{\theta}$, otherwise under the loss (2.1.5) every estimator of $\underline{\theta}$ is minimax having the supremum risk $+\infty$.

Now the general class of EB estimators of $\underline{\theta}$ is modified as

$$\widehat{\underline{\theta}}^{EB}(\widehat{\underline{V}}, \underline{X}) = (I_p - \widehat{\underline{V}}\,\underline{S}^{-1}\tau(\widehat{\underline{V}}, \underline{S}))\underline{X}, \tag{3.1.2}$$

where $\tau(\widehat{\underline{V}}, \underline{S})$ is a symmetric p×p matrix. Under the loss (3.1.1), the following theorem shows the relationship between the risks of $\underline{X}$ and a general EB estimator $\widehat{\underline{\theta}}^{EB}(\widehat{\underline{V}}, \underline{X})$ of the form (3.1.2). In this chapter we denote by $E_{\underline{\theta}, \underline{V}}$, the expectation conditional on $\underline{\theta}$ and $\underline{V}$,

$\tilde{E}$ the expectation over the marginal distribution of $\underline{\theta}$, and E the expectation over the joint distribution of $\underline{X}$, $\underline{\theta}$ and $\underline{Y}$. Also, let $R_{Q^*}(\underline{\theta}, \underline{e}(\hat{\underline{V}}, \underline{X})) = E_{\underline{\theta},\underline{Y}}[L_{Q^*}(\underline{\theta}, \underline{e}(\hat{\underline{V}}, \underline{X}))]$, where L_{Q^*} is defined in (3.1.1).

Theorem 3.1.1 Under the L_{Q^*} loss, $\hat{\underline{\theta}}^{EB}$ defined in (3.1.2) has risk

$$\begin{aligned} & R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB}) \\ = \; & R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta},\underline{Y}}\mathrm{tr}\{\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S})\underline{S}\,\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\,Q^*\hat{\underline{V}} - 4D^S[\underline{S}\,\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}]\hat{\underline{V}}\,Q^*\underline{V} \\ & - 2(K-p-1)\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\,Q^*\underline{V}\} \end{aligned} \tag{3.1.3}$$

Proof of Theorem 3.1.1 This proof is similar to the one in Theorem 2.2.1. Let $\hat{\underline{\theta}}_i^{EB}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - \hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S}))\underline{X}_i$, $i = 1, \ldots, K$. Then,

$$\begin{aligned} & EL_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB}) \\ = \; & E\{\sum_{i=1}^{K}(\underline{\theta}_i - \hat{\underline{\theta}}_i^{EB})^T Q^*(\underline{\theta}_i - \hat{\underline{\theta}}_i^{EB})\} \\ = \; & E\{\sum_{i=1}^{K}(\underline{\theta}_i - \hat{\underline{\theta}}_i^{B} + \hat{\underline{\theta}}_i^{B} - \hat{\underline{\theta}}_i^{EB})^T Q^*(\underline{\theta}_i - \hat{\underline{\theta}}_i^{B} + \hat{\underline{\theta}}_i^{B} - \hat{\underline{\theta}}_i^{EB})\} \\ = \; & EL_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{B}) + E\{\sum_{i=1}^{K}(\hat{\underline{\theta}}_i^{B} - \hat{\underline{\theta}}_i^{EB})^T Q^*(\hat{\underline{\theta}}_i^{B} - \hat{\underline{\theta}}_i^{EB})\} \\ = \; & EL_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{B}) + E\mathrm{tr}[(\hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S}) - \underline{V}\Sigma^{-1})\underline{S}(\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}} - \Sigma^{-1}\underline{V})Q^*], \end{aligned} \tag{3.1.4}$$

where $\hat{\underline{\theta}}^{B}$ is defined in (2.1.6).

$$\begin{aligned} & E\mathrm{tr}\{(\hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S}) - \underline{V}\Sigma^{-1})\underline{S}(\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}} - \Sigma^{-1}\underline{V})Q^*\} \\ = \; & E\mathrm{tr}\{\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S})\underline{S}\,\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\,Q^*\hat{\underline{V}} - 2\Sigma^{-1}\underline{S}\,\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\,Q^*\underline{V} \end{aligned}$$

$$+\underline{\Sigma}^{-1}\underline{S}\,\underline{\Sigma}^{-1}\underline{V}\,\underline{Q}^{*}\underline{V})]$$

$$= E\mathrm{tr}\{\underline{S}^{-1}\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}\,\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^{*}\hat{\underline{V}} - 2\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^{*}\underline{V}\}$$

$$+K\mathrm{tr}(\underline{\Sigma}^{-1}\underline{V}\,\underline{Q}^{*}\underline{V}), \tag{3.1.5}$$

since $\mathrm{tr}(\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}\,\underline{\Sigma}^{-1}\underline{V}\,\underline{Q}^{*}) = \mathrm{tr}(\underline{S}^{-1}\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}\,\underline{\Sigma}^{-1}\underline{V}\,\underline{Q}^{*}\hat{\underline{V}}) = \mathrm{tr}(\underline{S}^{-1}\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}\,\underline{\Sigma}^{-1}\underline{V}\,\underline{Q}^{*}\hat{\underline{V}})^{T} = \mathrm{tr}(\hat{\underline{V}}\underline{Q}^{*}\underline{V}\,\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}) = \mathrm{tr}(\underline{V}\,\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^{*}) = \mathrm{tr}(\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^{*}\underline{V})$ and $E(\underline{S}) = K\underline{\Sigma}$.

Substituting $\underline{Q} = \underline{Q}^{*}$ into (2.2.3) and using (3.1.5) one has

$$EL_{Q^{*}}(\underline{\theta},\widehat{\underline{\theta}}^{EB}) = K\mathrm{tr}(\underline{V}\,\underline{Q}^{*}) + E\mathrm{tr}\{\underline{S}^{-1}\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}\,\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^{*}\hat{\underline{V}}$$

$$-2\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^{*}\underline{V}\} \tag{3.1.6}$$

Note that $R_{Q^{*}}(\underline{\theta},\underline{X}) = E_{\underline{\theta},\underline{V}}L_{Q^{*}}(\underline{\theta},\underline{X}) = K\mathrm{tr}(\underline{V}\,\underline{Q}^{*}) = K\mathrm{tr}(\underline{Q})$ for all $\underline{\theta}$ and $\underline{V}$. Also applying the Wishart identity as given in Lemma 1.3.3, one obtains

$$EL_{Q^{*}}(\underline{\theta},\widehat{\underline{\theta}}^{EB})$$

$$= \tilde{E}R_{Q^{*}}(\underline{\theta},\widehat{\underline{\theta}}^{EB})$$

$$= \tilde{E}R_{Q^{*}}(\underline{\theta},\underline{X}) + \tilde{E}E_{\underline{\theta},\underline{V}}\mathrm{tr}\{\underline{S}^{-1}\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}\,\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^{*}\hat{\underline{V}}$$

$$-4\underline{D}^{S}[\underline{S}\,\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}]\hat{\underline{V}}\,\underline{Q}^{*}\underline{V} - 2(K-p-1)\underline{\tau}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^{*}\underline{V}\} \tag{3.1.7}$$

Note that $\hat{\underline{V}}$ is independent of $\underline{X}$ (and $\underline{S}$). Again, the conditional distribution of $\underline{S} = \sum_{i=1}^{K}\underline{X}_i\underline{X}_i^{T}$ given $\underline{\theta}$ depends on $\underline{\theta}$ only through $\sum_{i=1}^{K}\underline{\theta}_i\,\underline{\theta}_i^{T} = \underline{\theta}\,\underline{\theta}^{T} = \underline{\Theta}$(say) and $\hat{\underline{V}}$ is ancillary. Also the family of distributions of $\underline{\Theta}$ being Wishart $W(p, K, \underline{\Delta})$ is complete for $\underline{\Delta}$ (or $\underline{\Sigma}$). Thus, (3.1.3) follows immediately from (3.1.7).

The following assumptions of $\underline{V}$ with a complete sufficient statistic $\hat{\underline{V}}$ independent of $\underline{X}$ will be considered respectively in the next three sections.

(i) $\underline{V} = v\underline{G}$, $\underline{G}$ known symmetric p.d., and $v > 0$ unknown.

$$\hat{\underline{V}} = \hat{v}\underline{G},\ \hat{v} \sim v(n+2)^{-1}\chi_n^2, \tag{3.1.8}$$

where χ_n^2 is chi-square distribution with n degrees of freedom;

(ii) $\underline{V} = \underline{V}_d = \mathrm{diag}(v_1, \ldots, v_p)$, a diagonal matrix with ith diagonal element $v_i > 0$ and all v_i's unknown.

$$\hat{\underline{V}}_d = \mathrm{diag}(\hat{v}_1, \ldots, \hat{v}_p),\ \hat{v}_i \sim v_i(n+2)^{-1}\chi_n^2 \text{ independently}, \tag{3.1.9}$$

(iii) $\underline{V}$ p.d. and entirely unknown. With

$$\hat{\underline{V}} \sim (n+p+1)^{-1}W(p, n, \underline{V}),\ n > p+1 \tag{3.1.10}$$

The above three cases had been considered in Efron and Morris (1972b). They showed that the particular EB estimator $(\underline{I}_p - (K-p-1)\hat{\underline{V}}\underline{S}^{-1})\underline{X}$ dominated $\underline{X}$ under the loss given in (3.1.1) with $\underline{Q} = \underline{I}_p$.

Results similar to the ones in Chapter 2 are shown to hold (Section 3.2-3.4) for the cases (3.1.8)-(3.1.10). In Section 3.2-3.4, we first provide two classes of EB minimax estimators, namely $\hat{\underline{\theta}}^{EB(1)}$ and $\hat{\underline{\theta}}^{EB(2)}$, of the form (3.1.2) under the L_{Q^*} loss, where

$$\hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\underline{S}^{-1})\underline{X}, \tag{3.1.11}$$

where $r(\hat{\underline{V}}, \underline{S})$ (> 0) is a scalar, and

$$\hat{\underline{\theta}}^{EB(2)}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - \hat{\underline{V}}\underline{S}^{-1}\underline{\varphi}(\underline{S}))\underline{X}. \tag{3.1.12}$$

where $\underline{g}(\underline{S})$ is a symmetric p.d. matrix.

Next, as in Section 2.3, we show a class of estimators $\hat{\underline{\theta}}^{EB(3)}$ which dominate $\hat{\underline{\theta}}^{EB(1)}$ with certain $r(\hat{\underline{V}}, \underline{S})$ under the L_{Q^*} loss, where

$$\hat{\underline{\theta}}^{EB(3)}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\,\underline{S}^{-1} - \hat{\underline{V}}\underline{S}^{-1}\underline{f}(\underline{S}))\underline{X}, \tag{3.1.13}$$

and $\underline{f}(\underline{S})$ is symmetric as defined in Section 2.3. The risk difference of $\hat{\underline{\theta}}^{EB(3)}$ and $\hat{\underline{\theta}}^{EB(1)}$ is given in the following theorem.

Theorem 3.1.2 Under the L_{Q^*} loss with $\underline{f}(\underline{S}) = \underline{P}\,\underline{\Psi}(\underline{\lambda})\,\underline{P}^T$, $\underline{\Psi}(\underline{\lambda})$ is a diagonal matrix with ith diagonal element equal to $\Psi_i(\underline{\lambda}) > 0$, and assume $r(\hat{\underline{V}}, \underline{S})$ is differentiable with respect to s_{ij} for each i, j = 1, ..., p.

$$\begin{aligned} & R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(3)}(\hat{\underline{V}}, \underline{X})) - R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}, \underline{X})) \\ = \; & E_{\underline{\theta},\underline{V}}\, tr[\underline{P}[\underline{\Psi}^2(\underline{\lambda})\underline{\Delta}^{-1} + 2r(\hat{\underline{V}}, \underline{S})\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}]\underline{P}^T\hat{\underline{V}}\,\underline{Q}^*\hat{\underline{V}} \\ & -\underline{P}[4\underline{\Psi}^{(1)}(\underline{\lambda}) + 2(K-p-1)\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}]\underline{P}^T\hat{\underline{V}}\,\underline{Q}^*\underline{V}] \end{aligned} \tag{3.1.14}$$

Proof of Theorem 3.1.2 In view of (3.1.3), one gets

$$\begin{aligned} & R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(3)}(\hat{\underline{V}}, \underline{X})) - R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}, \underline{X})) \\ = \; & E_{\underline{\theta},\underline{V}}\, tr\{[\underline{S}^{-1}\underline{f}(\underline{S})\underline{S}\,\underline{f}(\underline{S})\underline{S}^{-1} + r(\hat{\underline{V}}, \underline{S})[\underline{S}^{-1}\underline{f}(\underline{S}) + \underline{f}(\underline{S})\underline{S}^{-1}]]\hat{\underline{V}}\,\underline{Q}^*\hat{\underline{V}} \\ & -[4\underline{D}^S[\underline{S}\,\underline{f}(\underline{S})\underline{S}^{-1}] + 2(K-p-1)\underline{f}(\underline{S})\underline{S}^{-1}]\hat{\underline{V}}\,\underline{Q}^*\underline{V}\} \end{aligned} \tag{3.1.15}$$

Hence (3.1.14) is obtained by using $\underline{f}(\underline{S}) = \underline{P}\,\underline{\Psi}(\underline{\lambda})\,\underline{P}^T$ and (2.2.18), and the proof is complete.

Finally, in Section 3.5, five different estimators of $\underline{\theta}$ are compared according to their simulated risks under the loss (3.1.1) and under three different assumptions of $\underline{V}$ given in (3.1.8)-(3.1.10). We continue the choice of p, K, $\underline{\theta}$, $\underline{V}$ and Q used in Section 2.4. As before, the estimator $\hat{\underline{\theta}}_{\alpha_0}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - \hat{\underline{V}}[(1-\alpha_0)\{(K-p-1)\underline{S}^{-1} + ((p-1)/tr(\underline{S}))\underline{I}_p\} + \alpha_0[(Kp-2)/tr(\underline{S})]\underline{I}_p])\underline{X}$ with suitable choice of α_0 has a definite edge over its rival estimators when $\underline{\theta} = \underline{0}$. However, for $\underline{\theta}$ far away from $\underline{0}$, among the estimators dominating $\underline{X}$, there is no clear-cut advantage of any particular estimator over the others. Thus, on an overall basis, we recommend the use of $\hat{\underline{\theta}}_{\alpha_0}(\hat{\underline{V}}, \underline{X})$ as an estimator of $\underline{\theta}$.

3.2 EB Minimax Estimators Under the L_{Q^*} loss and $\underline{V} = v\,\underline{G}$, Only v Unknown

In this section, the covariance matrix $\underline{V}$ is known up to a positive value v, that is, $\underline{V} = v\,\underline{G}$, where $\underline{G}$ is known symmetric p.d. Also $\hat{v} \sim v(n+2)^{-1}\chi^2_n$ is available. In the balanced one-way MANOVA model introduced in (2.1.2), suppose $\underline{V}_0 = v\,\underline{G}$ ($\underline{G}$ known symmetric p.d.), $n = Kp(m-1)$, and $\hat{\underline{V}} = m^{-1}\hat{v}\,\underline{G}$. The best multiple estimator of v is given by

$$\hat{v} = tr\Big(\sum_{i=1}^{K}\sum_{j=1}^{m}(\underline{X}_{ij}-\underline{X}_i)\underline{G}^{-1}(\underline{X}_{ij}-\underline{X}_i)^T\Big)/(n+2) \tag{3.2.1}$$

Note that conditional on $\underline{\theta}$, $\underline{X}_i$'s and $\hat{v}$ (or $\hat{\underline{V}}$) are independent, and the distribution of $\hat{v}$ does not depend on $\underline{\theta}$. Also marginally, $(\underline{S}, \hat{\underline{V}})$ is the complete sufficient statistic for $(\underline{\Sigma}, \underline{V})$, where $\underline{V} = m^{-1}\underline{V}_0$.

In this case, the estimators $\hat{\underline{\theta}}^{EB(1)}$ and $\hat{\underline{\theta}}^{EB(2)}$ defined in (3.1.11) and (3.1.12) are modified as

$$\hat{\underline{\theta}}^{EB(1)}(\hat{v}\,\underline{G}, \underline{X}) = (\underline{I}_p - r(\hat{v}\,\underline{G}, \underline{S})\hat{v}\,\underline{G}\,\underline{S}^{-1})\underline{X} \tag{3.2.2}$$

and

$$\hat{\underline{\theta}}^{EB(2)}(\hat{v}\underline{G}, \underline{X}) = (\underline{I}_p - \hat{v}\underline{G}\underline{S}^{-1}\underline{\tau}(\underline{S}))\underline{X}. \tag{3.2.3}$$

First the risk of $\hat{\underline{\theta}}^{EB(1)}(\hat{v}\underline{G}, \underline{X})$ is calculated as follows

Corollary 3.2.1 Under the L_{Q^*} loss defined in (3.1.1), $\hat{\underline{\theta}}^{EB(1)}(\hat{v}\underline{G}, \underline{X})$ has risk

$$\begin{aligned} & R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}) \\ = \; & R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta},v}\mathrm{tr}\{[\tau^2(\hat{v}\underline{G}, \underline{S}) - 2(K-p-1)\tau(\hat{v}\underline{G}, \underline{S})]\hat{v}\underline{S}^{-1/2}\underline{G}^{1/2}\underline{Q}\underline{G}^{1/2}\underline{S}^{-1/2} \\ & -4[\underline{D}^S\tau(\hat{v}\underline{G}, \underline{S})]\hat{v}\underline{G}^{1/2}\underline{Q}\underline{G}^{1/2} \\ & +(\tfrac{4}{n+2})\hat{v}^2\tau(\hat{v}\underline{G}, \underline{S})(\partial\tau(\hat{v}\underline{G}, \underline{S})/\partial\hat{v})\underline{S}^{-1/2}\underline{G}^{1/2}\underline{Q}\underline{G}^{1/2}\underline{S}^{-1/2} \end{aligned} \tag{3.2.4}$$

Proof of Corollary 3.2.1 Let $\underline{\tau}(\hat{\underline{V}}, \underline{S}) = \tau(\hat{\underline{V}}, \underline{S})\underline{I}_p$ and $\hat{\underline{V}} = \hat{v}\underline{G}$. It follows from (3.1.3) that

$$\begin{aligned} & R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}) \\ = \; & R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta},v}\mathrm{tr}\{[\tau^2(\hat{v}\underline{G}, \underline{S})\hat{v}^2\underline{S}^{-1} - 4[\underline{D}^S\tau(\hat{v}\underline{G}, \underline{S})]\hat{v}v \\ & -2(K-p-1)\tau(\hat{v}\underline{G}, \underline{S})\hat{v}v\underline{S}^{-1}]\underline{G}\underline{Q}^*\underline{G}\} \end{aligned} \tag{3.2.5}$$

Next, applying the formula for the chi-square distribution in Lemma 1.3.2.

$$\begin{aligned} & E_v[\tau^2(\hat{v}\underline{G}, \underline{S})\hat{v}^2 v^{-1}] \\ = \; & E_v[(\tfrac{n}{n+2})\tau^2(\hat{v}\underline{G}, \underline{S})\hat{v} + (\tfrac{2}{n+2})\hat{v}\{\tau^2(\hat{v}\underline{G}, \underline{S}) + \hat{v}(\partial\tau^2(\hat{v}\underline{G}, \underline{S})/\partial\hat{v})\}] \\ = \; & E_v[\tau^2(\hat{v}\underline{G}, \underline{S})\hat{v} + (\tfrac{4}{n+2})\hat{v}^2\tau(\hat{v}\underline{G}, \underline{S})(\partial\tau(\hat{v}\underline{G}, \underline{S})/\partial\hat{v})], \end{aligned} \tag{3.2.6}$$

where $\hat{v} \sim \frac{v}{n+2}\chi^2_n$ and τ is such that all expectations in (3.2.6) exist.

Combining (3.2.5) and (3.2.6), one gets

$$R_{Q^*}(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)})$$

$$= R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta},v}\mathrm{tr}\{[r^2(\hat{v}\underline{G}, \underline{S}) - 2(K-p-1)r(\hat{v}\underline{G}, \underline{S})]\hat{v}v\underline{S}^{-1}\underline{G}\,Q^*\underline{G}$$

$$-4[\underline{D}^S r(\hat{v}\underline{G}, \underline{S})]\hat{v}v\underline{G}\,Q^*\underline{G} + (\tfrac{4}{n+2})\hat{v}^2 v\, r(\hat{v}\underline{G}, \underline{S})(\partial r(\hat{v}\underline{G}, \underline{S})/\partial\hat{v})\underline{S}^{-1}\underline{G}\,Q^*\underline{G}, \quad (3.2.7)$$

Recall that $Q^* = \underline{V}^{-1/2}Q\underline{V}^{-1/2} = v^{-1}\underline{G}^{-1/2}Q\underline{G}^{-1/2}$, one has (3.2.4) from (3.2.7). The proof is complete.

The following theorem is an immediate result of Corollary 3.2.1

Theorem 3.2.1 $\widehat{\underline{\theta}}^{EB(1)}(\hat{v}\underline{G}, \underline{X})$ of the form (3.2.2) dominates $\underline{X}$ under the L_{Q^*} loss if

(i) $0 < r(\hat{v}\underline{G}, \underline{S}) < 2(K-p-1)$,

(ii) $r(\hat{v}\underline{G}, \underline{S})$ is non-increasing in $\hat{v}$,

(iii) $[\underline{D}^S r(\hat{v}\underline{G}, \underline{S})]$ is n.n.d..

Proof of Theorem 3.2.1 Since $\underline{S}^{-1/2}Q^{1/2}Q\underline{G}^{1/2}\underline{S}^{-1/2}$ and $\underline{G}^{1/2}Q\underline{G}^{1/2}$ are symmetric p.d. and $\hat{v} > 0$ with probability 1. It is easy to see that $R_{Q^*}(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)}) - R_{Q^*}(\underline{\theta}, \underline{X}) < 0$ under the assumptions (i)-(iii). This concludes the proof.

As in Corollary 2.2.2, we consider the special case $r(\hat{v}\underline{G}, \underline{S}) = r(F)$ where $F = \mathrm{tr}(\underline{S}\,\hat{\underline{V}}^{-1}) = \hat{v}^{-1}\,\mathrm{tr}(\underline{S}\,\underline{G}^{-1})$ in the next corollary

Corollary 3.2.2 Let $r(\hat{v}\underline{G}, \underline{S}) = r(F)$ where $F = \hat{v}^{-1}\,\mathrm{tr}(\underline{S}\,\underline{G}^{-1})$ and $r(F)$ is differentiable. Then under the loss (3.1.1), $\widehat{\underline{\theta}}^{EB(1)}(\hat{v}\underline{G}, \underline{X})$ dominates $\underline{X}$ if

(i) $0 < r(\hat{v}\underline{G}, \underline{S}) < 2(K-p-1)$,

(ii) $r(F)$ is non-decreasing in F

Proof of Corollary 3.2.2 Since $tr(\underline{S}\,\underline{G}^{-1}) > 0$, it is obvious that condition (ii) in Theorem 3.2.1 is satisfied. Also $\underline{D}^{\underline{S}} r(F) = r'(\hat{v}\underline{G})^{-1} \geq \underline{0}$ by using condition (ii) and $\underline{G}$ is p.d. We complete the proof

Remark 3.2.1 It can be shown that when $r(\hat{v}\underline{G}, \underline{S}) = a$ is a constant, the optimal choice is K-p-1 which leads to the matrix extension form of James Stein estimator $(\underline{I}_p - (K-p-1)\hat{v}\underline{G}\,\underline{S}^{-1})\underline{X}$. Obviously, all the conditions of Theorem 3.2.1 and Corollary 3.2.2 are met.

Remark 3.2.2 From the proof of Theorem 3.2.1 it follows that $\hat{\underline{\theta}}^{EB(1)}(\hat{v}\underline{G}, \underline{X})$ dominates $\underline{X}$ under the conditions (i)-(iii) of Theorem 3.2.1 and L_Q loss given in (2.1.5).

Before we show in Theorem 3.2.2 that some $\hat{\underline{\theta}}^{EB(2)}(\hat{v}\underline{G}, \underline{X})$ of the form (3.2.3) also dominate $\underline{X}$ under the L_{Q^*} loss, the risk of $\hat{\underline{\theta}}^{EB(2)}(\hat{v}\underline{G}, \underline{X})$ is derived here.

Corollary 3.2.3 The risk of EB estimators $\hat{\underline{\theta}}^{EB(2)}(\hat{v}\underline{G}, \underline{X})$ of the form (3.2.3) can be expressed as

$$R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(2)})$$

$$= R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta}, v}\, tr\{\tfrac{n}{n+2}\, v\, Q^{1/2}\underline{G}^{1/2}[\underline{S}^{-1}\underline{\varphi}(\underline{S})\,\underline{S}\,\underline{\varphi}(\underline{S})\underline{S}^{-1} - 4\underline{D}^{\underline{S}}[\underline{S}\,\underline{\varphi}(\underline{S})\underline{S}^{-1}]$$

$$2(K-p-1)\underline{\varphi}(\underline{S})\underline{S}^{-1}]\underline{G}^{1/2}Q^{1/2}\} \qquad (3.2.8)$$

Proof of Corollary 3.2.3 Replacing $\underline{g}(\hat{\underline{V}}, \underline{S}) = \underline{\varphi}(\underline{S})$ and $\hat{\underline{V}} = \hat{v}\underline{G}$ in (3.1.3), one gets

$$R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(2)})$$

$$= R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta}, v} tr\{[\underline{S}^{-1}\underline{g}(\underline{S})\,\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}\hat{v}^2 - 4\underline{D}^{\underline{S}}[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}]\hat{v} v$$

$$- 2(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}\hat{v} v]\underline{G}\,\underline{Q}^*\underline{G}\} \qquad (3.2.9)$$

Since $\hat{v}$ is independent of $\underline{X}$ and $\underline{S}$, and $\hat{v} \sim v(n+2)^{-1}\chi_n^2$, one has

$$R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(2)})$$

$$= R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta}, v} tr\{[\underline{S}^{-1}\underline{g}(\underline{S})\,\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1} - 4\underline{D}^{\underline{S}}\{\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}\}$$

$$- 2(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}]\frac{n}{n+2}v^2\underline{G}\,\underline{Q}^*\underline{G}\} \qquad (3.2.10)$$

The final expression (3.2.8) follows from (3.2.10) by the fact $\underline{Q}^* = \underline{V}^{-1/2}\underline{Q}\,\underline{V}^{-1/2} = v^{-1}\underline{G}^{-1/2}\underline{Q}\,\underline{G}^{-1/2}$. The proof is complete.

We continue to assume $\underline{g}(\underline{S}) = \underline{P}\,\underline{\Psi}(\underline{\lambda})\,\underline{P}^T$ as in Section 2.2. Then from (2.2.16)-(2.2.20), we have the following theorem.

Theorem 3.2.2 Suppose

(i) $0 < \Psi_i(\underline{\lambda}) < 2(K-p-1)$ for each $i = 1, \ldots, p$,

(ii) $\partial\Psi_i(\underline{\lambda})/\partial\lambda_i \geq 0$ for each $i = 1, \ldots, p$,

(iii) $\Psi_i(\underline{\lambda})$'s are similarly ordered with the λ_i's, that is,

$(\Psi_i(\underline{\lambda}) - \Psi_j(\underline{\lambda}))(\lambda_i - \lambda_j) \geq 0$ for every $1 \leq i, j \leq p$

Then $\hat{\underline{\theta}}^{EB(2)}$ of the form (3.2.3) dominates $\underline{X}$ under the L_{Q^*} loss.

Proof of Theorem 3.2.2 See (2.2.16)-(2.2.20) and the proof of Theorem 2.2.3

Remark 3.2.3 The $\underline{\Psi}(\underline{\lambda})$'s proposed in Example 2.2.2 2.2.6 and Remark 2.2.3 validate $\hat{\underline{\theta}}^{EB(2)}(\hat{v}\underline{G}, \underline{X}) = (\underline{I}_p - \hat{v}\underline{G}\,\underline{S}^{-1}\underline{g}(\underline{S}))\underline{X}$ as a better estimator than $\underline{X}$ under the L_{Q^*} loss.

Remark 3.2.4 In Theorem 3.2.2, it can be shown that the dominance of $\hat{\underline{\theta}}^{EB(2)}$ over $\underline{X}$ continues to hold under the same conditions and the L_Q loss.

Although we have shown estimators $\hat{\underline{\theta}}^{EB(1)}$ of the form (3 2.2) to be better than $\underline{X}$ in Theorem 3 2 1 there is room for improvement on them. Consider the following special form of the estimator $\hat{\underline{\theta}}^{EB(3)}$ of $\underline{\theta}$ defined in (3 1 13)

$$\hat{\underline{\theta}}^{EB(3)}(\hat{v}\underline{G}, \underline{X}) = (\underline{I}_p - r(\hat{v}\underline{G}, \underline{S})\hat{v}\underline{G}\,\underline{S}^{-1} - \hat{v}\underline{G}\,\underline{S}^{-1}\underline{f}(\underline{S}))\underline{X}, \tag{3.2.11}$$

and

$$\hat{\underline{\theta}}^{EB(1)}(\hat{v}\underline{G}, \underline{X}) = (\underline{I}_p - r(\hat{v}\underline{G}, \underline{S})\hat{v}\underline{G}\,\underline{S}^{-1})\underline{X}, \tag{3 2 12}$$

where $r(\hat{v}\underline{G}, \underline{S})$ is differentiable with respect to s_{ij} for each $i, j = 1, \ldots, p$, and $r(\hat{v}\underline{G}, \underline{S})$ is non-increasing in $\hat{v}$. Follow the assumption $\hat{\underline{V}} = \hat{v}\underline{G}$ with $\hat{v} \sim v(n+2)^{-1}\chi^2_n$ and Lemma (1.3.2), one has $E_v[\hat{v}^2 r(\hat{v}\underline{G}, \underline{S})] = vE_v[\hat{v}\, r(\hat{v}\underline{G}, \underline{S})] + 2v(n+2)^{-1}E_v[\hat{v}^2\partial r(\hat{v}\underline{G}, \underline{S})/\partial\hat{v}]$. Hence (3 1 14) yields

$$R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(3)}(\hat{v}\underline{G}, \underline{X})) - R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}(\hat{v}\underline{G}, \underline{X}))$$

$$= E_{\underline{\theta},v} tr\{\underline{P}[\underline{\Psi}^2(\underline{\lambda})\underline{\Delta}^{-1} + 2r(\hat{v}\underline{G}, \underline{S})\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}]\underline{P}^T\hat{v}^2\underline{G}\,\underline{Q}^*\underline{G}$$

$$- \underline{P}[4\underline{\Psi}^{(1)}(\underline{\lambda}) + 2(K-p-1)\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}]\underline{P}^T\hat{v}v\underline{G}\,\underline{Q}^*\underline{G}\}$$

$$= E_{\underline{\theta},v} tr\{\hat{v}\underline{Q}^{1/2}\underline{G}^{1/2}\underline{P}[\underline{\Psi}^2(\underline{\lambda})\underline{\Delta}^{-1} - 2(K-p-1-r(\hat{v}\underline{G}, \underline{S}))\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}$$

$$-4\underline{\Psi}^{(1)}(\underline{\lambda})+(\tfrac{4}{n+2})\hat{v}(\partial r(\hat{v}\underline{G},\underline{S})/\partial\hat{v})]\underline{P}^T\underline{G}^{1/2}\underline{Q}^{1/2}]$$

$$\le E_{\underline{\theta},v}tr[\hat{v}\underline{Q}^{1/2}\underline{G}^{1/2}\underline{P}[\underline{\Psi}^2(\underline{\lambda})\underline{\Delta}^{-1}-2(K-p-1-r(\hat{v}\underline{G},\underline{S}))\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}$$

$$-4\underline{\Psi}^{(1)}(\underline{\lambda})]\underline{P}^T\underline{G}^{1/2}\underline{Q}^{1/2}] < 0, \tag{3.2.13}$$

for all $\underline{\theta}$ and v if $\Delta_{1i}(\underline{\lambda}) < 0$ for each $i = 1, \ldots, p$, where

$$\Delta_{1i}(\underline{\lambda}) = \Psi_i^2(\underline{\lambda})\lambda_i^{-1}-2(K-p-1-r(\hat{v}\underline{G},\underline{S}))\Psi_i(\underline{\lambda})\lambda_i^{-1}-2\sum_{j(\neq i)}\{\Psi_i(\underline{\lambda})-\Psi_j(\underline{\lambda})\}/(\lambda_i-\lambda_j)$$

$$-4[\partial\Psi_i(\underline{\lambda})/\partial\lambda_i] \tag{3.2.14}$$

Hence we have a generalization of Theorem 2.3.1

Theorem 3.2.3 Let $0 < r(\hat{v}\underline{G},\underline{S}) < (K-2)$, $r(\hat{v}\underline{G},\underline{S})$ is differentiable with respect to s_{ij} for each $i, j = 1, \ldots, p$. Suppose $\Psi_i(\underline{\lambda}) = \lambda_i\tau(U)/(\sum_{j=1}^{p}\lambda_j)$, $i = 1, \ldots, p$, where U is a function of the λ_i's. Suppose that

(i) $r(\hat{v}\underline{G},\underline{S})$ is non-increasing in $\hat{v}$.

(ii) $U(\underline{\lambda})$ is non-decreasing in λ_i for each $i = 1, \ldots, p$,

(iii) $\tau(U)$ is non-decreasing in U,

(iv) $0 < \tau(U) < 2(K-r(\hat{v}\underline{G},\underline{S})-2)$

Then $\hat{\underline{\theta}}^{EB(3)}(\hat{v}\underline{G},\underline{X})$ dominates $\hat{\underline{\theta}}^{EB(1)}(\hat{v}\underline{G},\underline{X})$ under the L_{Q^*} loss

Proof of Theorem 3.2.3 Same as the proof of Theorem 2.3.1 with the obvious change of $r(\underline{Y},\underline{S})$ into $r(\hat{v}\underline{G},\underline{S})$

Similar to Examples 2 3 1 and 2 3 2 and Remark 2 3 1, one has the following remark

Remark 3.2.5 When $r(\hat{v}\underline{G}, \underline{S}) = a$ is constant, $(I_p - a\hat{v}\underline{G}\underline{S}^{-1}[t(U)/tr(\underline{S})]\hat{v}\underline{G})\underline{X}$ dominates $(I_p - a\hat{v}\underline{G}\underline{S}^{-1})\underline{X}$ whenever $0 < a < (K-2)$ and $0 < t(U) < 2(K-a-2)$ For the particular choice of $a = K-p-1$, $0 < t(U) < 2(p-1)$. If $t(U) = b$ is constant, and $a = K-p-1$, it follows that the optimal choice of b is $p-1$ Note that when U equals the geometric, arithmetic or harmonic mean of the eigen values of $\underline{S}$, then the condition (ii) of Theorem 3 2.3 holds.

Remark 3.2.6 In Theorem 3 2 3, $\hat{\underline{\theta}}^{EB(3)}(\hat{v}\underline{G}, \underline{X})$ also dominates $\hat{\underline{\theta}}^{EB(1)}(\hat{v}\underline{G}, \underline{X})$ under the L_Q loss.

3.3 EB Minimax Estimators Under the L_{Q^*} loss and $\underline{V} = \underline{V}_d$ an Unknown Diagonal Matrix

A more general assumption regarding the covariance matrix of $\underline{X}_i$'s is that $\underline{V} = \underline{V}_d = \mathrm{diag}(v_1, \ldots, v_p)$ where $\mathrm{diag}(v_1, \ldots, v_p)$ is a diagonal matrix with ith diagonal element equal to v_i Assume the estimator $\hat{\underline{V}}_d = \mathrm{diag}(\hat{v}_1, \ldots, \hat{v}_p)$ is available, where $\hat{v}_i$ is independently distributed as $v_i(n+2)^{-1}\chi_n^2$ Also $\underline{X}_i$ and $\hat{v}_i$ are independent. For example, suppose $\underline{V}_0 = \underline{V}_d = \mathrm{diag}(v_1, \ldots, v_p)$, an unknown diagonal matrix with ith diagonal element $v_i > 0$, in the balanced one way MANOVA model given in (2.1.2) Let $n = K(m-1)$, $\hat{\underline{V}} = \hat{\underline{V}}_d = \mathrm{diag}(\hat{v}_1, \ldots, \hat{v}_p)$, where $\hat{v}_i$ is $[m(n+2)]^{-1}$ times the ith diagonal element of $(\sum_{i=1}^{K}\sum_{j=1}^{m}(\underline{X}_{ij}-\underline{X}_i)(\underline{X}_{ij}-\underline{X}_i)^T)$ Then $\hat{v}_i \sim v_i[m(n+2)]^{-1}\chi_n^2$ independently Also conditional on $\underline{\theta}$, $\underline{X}_i$'s and $\hat{\underline{V}}$ are independent, and the distribution of $\hat{\underline{V}}$ does not depend on $\underline{\theta}$. Marginally, $(\underline{S}, \hat{\underline{V}})$ is the complete sufficient statistic for $(\underline{\Sigma}, \underline{V})$ It is desired to estimate $\underline{\theta}$ under the quadratic loss (3.1.1) with $Q^* = \underline{V}_d^{-1/2} Q \underline{V}_d^{-1/2}$ and $\underline{V}_d = \underline{V}_d^{1/2}\underline{V}_d^{1/2}$. As before, we construct two classes of estimators, $\widehat{\underline{\theta}}^{EB(1)}$ and $\widehat{\underline{\theta}}^{EB(2)}$ purported to dominate $\underline{X}$ in their risks. Let

$$\widehat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}_d, \underline{X}) = (\underline{I}_p - r(\hat{\underline{V}}_d, \underline{S})\hat{\underline{V}}_d \underline{S}^{-1})\underline{X}, \qquad (3.3.1)$$

where $r(\hat{\underline{V}}_d, \underline{S})\ (>0)$, and

$$\widehat{\underline{\theta}}^{EB(2)}(\hat{\underline{V}}_d, \underline{X}) = (\underline{I}_p - \hat{\underline{V}}_d \underline{S}^{-1}\underline{g}(\underline{S}))\underline{X}, \qquad (3.3.2)$$

where $\underline{g}(\underline{S})$ is a symmetric p.d. matrix.

The following corollary provides the risk for $\widehat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}_d, \underline{X})$ under the loss (3.1.1) with $Q^* = \underline{V}_d^{-1/2} Q \underline{V}_d^{-1/2}$.

Corollary 3.3.1 $\widehat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}_d, \underline{X})$ as defined in (3.3.1) has risk

$$\begin{aligned} & R_{Q^*}(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)}) \\ = \; & R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta}, \underline{V}_d} \mathrm{tr}\{(\tfrac{n}{n+2}) r^2(\hat{\underline{V}}_d, \underline{S})\underline{S}^{-1}\hat{\underline{V}}_d Q^* \underline{V}_d \\ & + (\tfrac{2}{n+2}) r^2(\hat{\underline{V}}_d, \underline{S})\underline{S}^{-1}\hat{\underline{V}}_d \mathrm{diag}(Q^*)\underline{V}_d - 2(K-p-1) r(\hat{\underline{V}}_d, \underline{S})\underline{S}^{-1}\hat{\underline{V}}_d Q^* \underline{V}_d \\ & + (\tfrac{4}{n+2}) r(\hat{\underline{V}}_d, \underline{S})\underline{S}^{-1}\underline{V}_d [\underline{\nabla} r(\hat{\underline{V}}_d, \underline{S})]\hat{\underline{V}}_d Q^* \hat{\underline{V}}_d - 4[\underline{D}^S r(\hat{\underline{V}}_d, \underline{S})]\hat{\underline{V}}_d Q^* \underline{V}_d\}, \end{aligned} \qquad (3.3.3)$$

where $\mathrm{diag}(Q^*) = \mathrm{diag}(q_{11}^*, \ldots, q_{pp}^*)$, $Q^* = ((q_{ij}^*))$ and $\underline{\nabla} r(\hat{\underline{V}}_d, \underline{S})$ is a diagonal matrix with ith diagonal element equal to $\partial r(\hat{\underline{V}}_d, \underline{S})/\partial \hat{v}_i$.

Proof of Corollary 3.3.1 Suppose $\underline{r}(\hat{\underline{V}}, \underline{S}) = r(\hat{\underline{V}}, \underline{S})\underline{I}_p$, $\underline{V} = \underline{V}_d$ and $\hat{\underline{V}} = \hat{\underline{V}}_d$. It follows from (3.1.3) that

$$R_{Q^*}(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)})$$

$$= R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta}, \underline{V}_d} \mathrm{tr}\{ r^2(\underline{\hat{V}}_d, \underline{S}) \underline{S}^{-1} \underline{\hat{V}}_d Q^* \underline{\hat{V}}_d - 4[D^S r(\underline{\hat{V}}_d, \underline{S})] \underline{\hat{V}}_d Q^* \underline{V}_d$$

$$-2(K-p-1) r(\underline{\hat{V}}_d, \underline{S}) \underline{S}^{-1} \underline{\hat{V}}_d Q^* \underline{V}_d\} \qquad (3.3.4)$$

Let $E_{\underline{V}_d}[r^2(\underline{\hat{V}}_d, \underline{S}) \underline{V}_d^{-1} \underline{\hat{V}}_d Q^* \underline{\hat{V}}_d] = E_{\underline{V}_d}[W] = E_{\underline{V}_d}[((w_{ij}))]$, where $w_{ij} = r^2(\underline{\hat{V}}_d, \underline{S}) v_i^{-1} \hat{v}_i q_{ij}^* \hat{v}_j$

Since $\hat{v}_i$ and $\underline{X}_i$'s are independent and $\hat{v}_i \sim v_i(n+2)^{-1}\chi^2_n$, using Lemma 1.3.2, one gets

$$E_{v_i}[w_{ii}]$$

$$= E_{v_i}[r^2(\underline{\hat{V}}_d, \underline{S}) v_i^{-1} \hat{v}_i^2 q_{ii}^*]$$

$$= E_{v_i}[\{r^2(\underline{\hat{V}}_d, \underline{S})\hat{v}_i + (\tfrac{4}{n+2}) \hat{v}_i^2 r(\underline{\hat{V}}_d, \underline{S})(\partial r(\underline{\hat{V}}_d, \underline{S})/\partial \hat{v}_i)\} q_{ii}^*] \qquad (3.3.5)$$

and

$$E_{v_i}[w_{ij}]$$

$$= E_{v_i}[r^2(\underline{\hat{V}}_d, \underline{S}) v_i^{-1} \hat{v}_i q_{ij}^* \hat{v}_j]$$

$$= E_{v_i}[\{(\tfrac{n}{n+2}) r^2(\underline{\hat{V}}_d, \underline{S}) + (\tfrac{4}{n+2}) \hat{v}_i r(\underline{\hat{V}}_d, \underline{S})(\partial r(\underline{\hat{V}}_d, \underline{S})/\partial \hat{v}_i)\} q_{ij}^* \hat{v}_j] \qquad (3.3.6)$$

Combining (3.3.5) and (3.3.6) to get

$$E_{\underline{\theta}, \underline{V}_d} \mathrm{tr}[r^2(\underline{\hat{V}}_d, \underline{S}) \underline{S}^{-1} \underline{\hat{V}}_d Q^* \underline{\hat{V}}_d]$$

$$= E_{\underline{\theta}, \underline{V}_d} \mathrm{tr}[(\tfrac{n}{n+2}) r^2(\underline{\hat{V}}_d, \underline{S}) \underline{S}^{-1} \underline{V}_d Q^* \underline{\hat{V}}_d + (\tfrac{2}{n+2}) r^2(\underline{\hat{V}}_d, \underline{S}) \underline{S}^{-1} \underline{V}_d \mathrm{diag}(Q^*) \underline{\hat{V}}_d$$

$$+ (\tfrac{4}{n+2}) r(\underline{\hat{V}}_d, \underline{S}) \underline{S}^{-1} \underline{V}_d [\nabla r(\underline{\hat{V}}_d, \underline{S})] \underline{\hat{V}}_d Q^* \underline{\hat{V}}_d]$$

$$= E_{\underline{\theta}, \underline{V}_d} \mathrm{tr}[(\tfrac{n}{n+2}) r^2(\underline{\hat{V}}_d, \underline{S}) \underline{S}^{-1} \underline{\hat{V}}_d Q^* \underline{V}_d + (\tfrac{2}{n+2}) r^2(\underline{\hat{V}}_d, \underline{S}) \underline{S}^{-1} \underline{\hat{V}}_d \mathrm{diag}(Q^*) \underline{V}_d$$

$$+\left(\tfrac{4}{n+2}\right)r(\hat{\underline{V}}_d, \underline{S})\underline{S}^{-1}\underline{V}_d[\underline{\nabla}r(\hat{\underline{V}}_d, \underline{S})]\hat{\underline{V}}_d Q^* \hat{\underline{V}}_d] \qquad (3.3.7)$$

Substituting (3.3.7) into (3.3.4) gives (3.3.3). Corollary 3.3.1 is proved.

Assume the loss function (3.1.1) with $\underline{Q} = \underline{Q}_d = \mathrm{diag}(q_{11}, \ldots, q_{pp})$. Recall that $\underline{Q} = ((q_{ij}))$. The dominance conditions of $\widehat{\underline{\theta}}^{EB(1)}$ over $\underline{X}$ are provided in the next theorem.

Theorem 3.3.1 Suppose $\underline{Q} = \underline{Q}_d = \mathrm{diag}(q_{11}, \ldots, q_{pp})$ in (3.1.1) and

(i) $0 < r(\hat{\underline{V}}_d, \underline{S}) < 2(K-p-1)$

(ii) $r(\hat{\underline{V}}_d, \underline{S})$ is non-increasing in $\hat{v}_i$ for each $i = 1, \ldots, p$,

(iii) $r(\hat{\underline{V}}_d, \underline{S})$ is non-decreasing in s_{ii} for each $i = 1, \ldots, p$,

where $\underline{S} = ((s_{ij}))$. Then $R_{Q^*}(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)}) < R_{Q^*}(\underline{\theta}, \underline{X})$.

Proof of Theorem 3.3.1 From (3.3.3), and using the assumption that $Q = \mathrm{diag}(q_{11}, \ldots, q_{pp})$ is a diagonal matrix,

$$\begin{aligned}
& R_{Q^*}(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)}) \\
= \;& R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta}}\underline{V}_d\mathrm{tr}\{[r^2(\hat{\underline{V}}_d, \underline{S}) - 2(K-p-1)r(\hat{\underline{V}}_d, \underline{S})]\underline{S}^{-1}\hat{\underline{V}}_d Q^* \underline{V}_d \\
& + \tfrac{4}{n+2}r(\hat{\underline{V}}_d, \underline{S})\underline{S}^{-1}\underline{V}_d[\underline{\nabla}r(\hat{\underline{V}}_d, \underline{S})]\hat{\underline{V}}_d Q^* \hat{\underline{V}}_d - 4[\underline{D}^S r(\hat{\underline{V}}_d, \underline{S})]\hat{\underline{V}}_d Q^* \underline{V}_d] \\
\le \;& R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta}}\underline{V}_d\mathrm{tr}\{[r^2(\hat{\underline{V}}_d, \underline{S}) - 2(K-p-1)r(\hat{\underline{V}}_d, \underline{S})]\underline{S}^{-1}\hat{\underline{V}}_d Q^* \underline{V}_d] \qquad (3.3.8)
\end{aligned}$$

The inequality in (3.3.8) uses diagonality of $\underline{V}_d$, $\hat{\underline{V}}_d$ and Q^* and conditions (ii) and (iii). The theorem now follows from (3.3.8) and condition (i).

Consider the particular choice of $r(\hat{\underline{V}}_d, \underline{S}) = r(F)$, where $F = \mathrm{tr}(\underline{S}\,\hat{\underline{V}}_d^{-1})$

Corollary 3.3.2 Let $r(\underline{\hat{V}}_d, \underline{S}) = r(F)$, where $F = \mathrm{tr}(\underline{S}\,\underline{\hat{V}}_d^{-1})$. Then $\underline{\hat{\theta}}^{EB(1)}(\underline{\hat{V}}_d, \underline{S}) = (\underline{I}_p - r(F)\underline{\hat{V}}_d \underline{S}^{-1})\underline{X}$ dominates $\underline{X}$ under the L_{Q^*} loss with $Q = Q_d$ a diagonal matrix if

(i) $0 < r(F) < 2(K-p-1)$,

(ii) $r(F)$ is non-decreasing in F

Proof of Corollary 3.3.2 Since $\partial r(F)/\partial \hat{v}_i = r'(\partial F/\partial \hat{v}_i) \le 0$ and $\partial F/\partial s_{ii} = r'(\partial F/\partial s_{ii}) \ge 0$ conditions (ii) and (iii) of Theorem 3.3.1 are satisfied. The corollary follows now by using condition (i)

Remark 3.3.1 The optimal choice of $r(\underline{\hat{V}}_d, \underline{S}) = a$ (constant) in Theorem 3.3.1 and Corollary 3.3.2 is K-p-1 which leads to another matrix extension of James-Stein estimator $(\underline{I}_p - (K-p-1)\underline{\hat{V}}_d \underline{S}^{-1})\underline{X}$.

Remark 3.3.2 By examining the proof of Theorem 3.3.1, under the same conditions and L_Q loss with $Q = Q_d$, the risk of $\underline{\hat{\theta}}^{EB(1)}$ is smaller than that of $\underline{X}$.

The risk $R_{Q^*}(\underline{\theta}, \underline{\hat{\theta}}^{EB(2)})$ of $\underline{\hat{\theta}}^{EB(2)}$ defined in (3.3.2) is provided here.

Corollary 3.3.3 Under the L_{Q^*} loss, $\underline{\hat{\theta}}^{EB(2)}(\underline{\hat{V}}_d, \underline{X}) = (\underline{I}_p - \underline{\hat{V}}_d \underline{S}^{-1}\underline{\phi}(\underline{S}))\underline{X}$ has risk

$$R_{Q^*}(\underline{\theta}, \underline{\hat{\theta}}^{EB(2)})$$

$$= R_{Q^*}(\underline{\theta}, \underline{X}) + \frac{n}{(n+2)^2} E_{\underline{\theta}, \underline{V}_d} \mathrm{tr}\{[n\underline{S}^{-1}\underline{\phi}(\underline{S})\underline{S}\,\underline{\phi}(\underline{S})\underline{S}^{-1} - 4(n+2)\underline{D}^{S}[\underline{S}\,\underline{\phi}(\underline{S})\underline{S}^{-1}]$$

$$- 2(n+2)(K-p-1)\underline{\phi}(\underline{S})\underline{S}^{-1}]\underline{V}_d^{1/2}\underline{Q}\underline{V}_d^{1/2} + 2\underline{S}^{-1}\underline{\phi}(\underline{S})\underline{S}\,\underline{\phi}(\underline{S})\underline{S}^{-1}\underline{V}_d^{1/2}\underline{Q}_d\underline{V}_d^{1/2}\} \qquad (3.3.9)$$

Proof of Corollary 3.3.3 Replacing $\underline{\tau}(\hat{\underline{V}}, \underline{S}) = \underline{g}(\underline{S})$ and $\hat{\underline{V}} = \hat{\underline{V}}_d$ in (3.1.3), one gets the expression

$$R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(2)})$$

$$= R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta}, \underline{V}_d} tr\{\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}\hat{\underline{V}}_d Q^* \hat{\underline{V}}_d - 4D^S[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}]\hat{\underline{V}}_d Q^* \underline{V}_d$$

$$-2(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}\hat{\underline{V}}_d Q^* \underline{V}_d\} \qquad (3.3.10)$$

Note that

$$E_{\underline{V}_d}[\hat{\underline{V}}_d Q^* \hat{\underline{V}}_d] = E_{\underline{V}_d}[((\hat{v}_i q^*_{ij} \hat{v}_j))]$$

$$= (\tfrac{n}{n+2})^2 \underline{V}_d Q^* \underline{V}_d + [\tfrac{2n}{(n+2)^2}]\, \underline{V}_d \operatorname{diag}(Q^*) \underline{V}_d \qquad (3.3.11)$$

and

$$E_{\underline{V}_d}[\hat{\underline{V}}_d Q^* \underline{V}_d)] = (\tfrac{n}{n+2})\, \underline{V}_d Q^* \underline{V}_d \qquad (3.3.12)$$

By using (3.3.11) and (3.3.12), (3.3.10) can be expressed as

$$R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(2)})$$

$$= R_{Q^*}(\underline{\theta}, \underline{X}) + \frac{n}{(n+2)^2} E_{\underline{\theta}, \underline{V}_d} tr\{[n\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1} - 4(n+2)D^S[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}]$$

$$- 2(n+2)(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}]\underline{V}_d Q^* \underline{V}_d + 2\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}\underline{V}_d \operatorname{diag}(Q^*)\underline{V}_d\} \qquad (3.3.13)$$

Next, (3.3.9) follows (3.3.13) by the fact $Q^* = \underline{V}_d^{-1/2} Q \underline{V}_d^{-1/2}$ and $\operatorname{diag}(Q^*) = \underline{V}_d^{-1/2} Q_d \underline{V}_d^{-1/2}$. The theorem is complete.

Notice that $Q_d = \operatorname{diag}(q_{11}, \ldots, q_{pp}) \le (\max q_{ii}) I_p \le (\max q_{ii})/(\min \eta_i) Q$, where $(\max q_{ii}) \ge q_{ii}$ and $(\min \eta_i) \le \eta_i$ for each $i = 1, \ldots, p$, and η_i's are eigen values of Q. Also

$\underline{S}^{-1}\underline{\varphi}(\underline{S})\underline{S}\,\underline{\varphi}(\underline{S})\underline{S}^{-1}$ is symmetric p.d., one has $\mathrm{tr}[\underline{S}^{-1}\underline{\varphi}(\underline{S})\underline{S}\,\underline{\varphi}(\underline{S})\underline{S}^{-1}\underline{V}_d^{1/2}\underline{Q}\,\underline{V}_d^{1/2}] \le \mathrm{tr}[\underline{S}^{-1}\underline{\varphi}(\underline{S})\underline{S}\,\underline{\varphi}(\underline{S})\underline{S}^{-1}\underline{V}_d^{1/2}(\max q_{ii})/(\min \eta_i)\underline{Q}\,\underline{V}_d^{1/2}]$.

It follows from (3.3.9) that

$$
\begin{aligned}
& R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(2)}) \\
\le\ & R_{Q^*}(\underline{\theta}, \underline{X}) + \left(\frac{n}{(n+2)^2}\right) E_{\underline{\theta}, \underline{V}_d}\mathrm{tr}\{[(n+2(\max q_{ii})/(\min \eta_i))\underline{S}^{-1}\underline{\varphi}(\underline{S})\underline{S}\,\underline{\varphi}(\underline{S})\underline{S}^{-1} \\
& -4(n+2)\underline{D}^{S}[\underline{S}\,\underline{\varphi}(\underline{S})\underline{S}^{-1}]-2(n+2)(K-p-1)\underline{\varphi}(\underline{S})\underline{S}^{-1}]\underline{V}_d^{1/2}\underline{Q}\,\underline{V}_d^{1/2}\} \\
=\ & R_{Q^*}(\underline{\theta}, \underline{X}) + \left(\frac{n}{n+2}\right) r_0^{-1} E_{\underline{\theta}, \underline{V}_d}\mathrm{tr}\{[\underline{S}^{-1}\underline{\varphi}(\underline{S})\underline{S}\,\underline{\varphi}(\underline{S})\underline{S}^{-1}-4r_0\underline{D}^{S}[\underline{S}\,\underline{\varphi}(\underline{S})\underline{S}^{-1}] \\
& -2r_0(K-p-1)\underline{\varphi}(\underline{S})\underline{S}^{-1}]\underline{V}_d^{1/2}\underline{Q}\,\underline{V}_d^{1/2}\}, \qquad (3.3.14)
\end{aligned}
$$

where $r_0 = (n+2)/[n+2(\max q_{ii})/(\min \eta_i)]$.

The inequality in (3.3.14) becomes an equality if $\underline{Q}$ is a diagonal matrix which yields $r_0 = 1$.

Let $\underline{S} = \underline{P}\,\underline{\Lambda}\,\underline{P}^T$ and $\underline{\varphi}(\underline{S}) = \underline{P}\,\underline{\Psi}(\underline{\lambda})\,\underline{P}^T$ as before. Using the same argument described in Section 2.2, we prove the next theorem which generalizes Zheng's (1986) result.

Theorem 3.3.2 $\hat{\underline{\theta}}^{EB(2)}$ given in (3.3.2) has risk smaller than $\underline{X}$ under the L_{Q^*} loss if

(i) $0 < \Psi_i(\underline{\lambda}) < 2r_0(K-p-1)$ for each $i = 1, \ldots, p$,

(ii) $\partial\Psi_i(\underline{\lambda})/\partial\lambda_i \ge 0$ for each $i = 1, \ldots, p$,

(iii) $\Psi_i(\underline{\lambda})$'s are similarly ordered with the λ_i's, that is,

$(\Psi_i(\underline{\lambda})-\Psi_j(\underline{\lambda}))(\lambda_i-\lambda_j) \ge 0$ for every $1 \le i, j \le p$

Proof of Theorem 3.3.2 It suffices to show

$$u_i(\lambda_i) = \Psi_i^2(\lambda_i)\lambda_i^{-1} - 4r_0\Psi_i^{(1)}(\lambda_i) - 2r_0(K-p-1)\Psi_i(\lambda_i)\lambda_i^{-1} < 0 \qquad (3.3.15)$$

for each $i = 1, \ldots, p$, where $\Psi_i^{(1)}(\lambda_i)$ is defined in (2.2.19) and is positive under the assumptions (ii) and (iii). Hence $u_i(\lambda_i) \le \lambda_i^{-1}\Psi_i(\lambda_i)[\Psi_i(\lambda_i) - 2r_0(K-p-1)] < 0$, by using assumption (i). The proof of the theorem is complete.

Remark 3.3.3 The range for $\Psi_i(\lambda_i)$ in Theorem 3.3.2 is smaller than the one in Theorem 2.2.3 which reflects the price for estimating the unknown v_i's from $\hat{v}_i$'s. Let $\underline{\Delta}^{-1}\underline{\Psi}(\underline{\lambda})$ be a diagonal matrix with ith diagonal element equal to (2.2.22). Then if $a \ge 0$, $b \ge 0$ and $0 < a+b < 2r_0(K-p-1)$, $[\underline{I}_p - \underline{\hat{V}}_d(a\underline{S}^{-1} + (b/tr(\underline{S}))\underline{I}_p)]\underline{X}$ and $[\underline{I}_p - \underline{\hat{V}}_d[a\underline{S}^{-1} + (b/tr(\underline{S}^2))\underline{S}]]\underline{X}$ dominate $\underline{X}$ for every L_{Q^*} loss. For estimators not meeting the conditions of Theorem 3.3.2, such as $(\underline{I}_p - \underline{\hat{V}}_d\underline{S}^{-1}\underline{g}(\underline{S}))\underline{X}$ with $\underline{\Psi}(\underline{\lambda})$ defined in (2.2.24) and (2.2.27) and (2.2.30) dominate $\underline{X}$ if (a) $0 < b < 2r_0(K-p-1)$, (b) $c > 0$, and (c) $0 \le d \le p(K-p-1)/2$, and if (a) $0 < b < 2r_0(K-p-1)$, (b) $c > 0$, and (c) $0 \le d \le (K-p-1)/2$, respectively. Also $(\underline{I}_p - \underline{\hat{V}}_d[\underline{S}^{-1}\underline{g}(\underline{S})]_\alpha)\underline{X}$ dominates $\underline{X}$ under L_{Q^*} loss if $\alpha \in [0, c_0^{-1}(K-p-1)(p-1)\{1+[1+c_0(p-1)^{-2}(K-p-1)^{-2}((K-p-1)^2 + (p-1)^2)]^{1/2}\}]$, $c_0 = [(K-2) - r_0^{-1}(Kp-2)]^2$, where

$$\{\underline{S}^{-1}\underline{g}(\underline{S})\}_\alpha = (1-\alpha)[(K-p-1)\underline{S}^{-1} + ((p-1)/tr(\underline{S}))\underline{I}_p] + \alpha[((Kp-2)/tr(\underline{S}))\underline{I}_p] \qquad (3.3.16)$$

and the optimal value of α is $c_0^{-1}(p-1)(K-p-1)$.

Remark 3.3.4 It can be shown that estimators of the form $\hat{\underline{\theta}}^{EB(2)}$ given in (3.3.2) also dominate $\underline{X}$ under the conditions (i)-(iii) in Theorem 3.3.2 and L_Q loss given in (2.1.5).

Next, we consider estimators of the form (3.1.13) with $\underline{\hat{V}} = \underline{\hat{V}}_d$

$$\underline{\hat{\theta}}^{EB(3)}(\hat{\underline{V}}, \underline{X}) = (I_p - r(\hat{\underline{V}}_d, \underline{S})\hat{\underline{V}}_d \underline{S}^{-1}\hat{\underline{V}}_d \underline{S}^{-1} f(\underline{S}))\underline{X} \qquad (3\,3\,17)$$

From Theorem 3 1.2 it follows that

$$R_{Q^*}(\underline{\theta}, \underline{\hat{\theta}}^{EB(3)}(\hat{\underline{V}}_d, \underline{X})) - R_{Q^*}(\underline{\theta}, \underline{\hat{\theta}}^{EB(1)}(\hat{\underline{V}}_d, \underline{X})]$$

$$= E_{\underline{\theta}, \underline{V}_d} tr\{[\underline{E}(\underline{\Psi}^2(\underline{L})\Delta^{-1} + 2r(\hat{\underline{V}}_d, \underline{S})\underline{\Psi}(\underline{L})\Delta^{-1}]\underline{P}^T\hat{\underline{V}}_d \underline{Q}^* \hat{\underline{V}}_d$$

$$-\underline{P}[4\underline{\Psi}^{(1)}(\underline{L}) + 2(K-p-1)\underline{\Psi}(\underline{L})\Delta^{-1}]\underline{P}^T\hat{\underline{V}}_d \underline{Q}^* \underline{V}_d\}, \qquad (3\,3\,18)$$

where $r(\hat{\underline{V}}_d, \underline{S})$ is differentiable with respect to s_{ij} for each $i, j = 1, \ldots, p$, and $r(\hat{\underline{V}}_d, \underline{S})$ is non increasing in $\hat{v}_i$ for each $i = 1, \ldots, p$

Recall that $\hat{v}_i$'s and $\underline{X}_i$'s are independent. Using the chi square identity in Lemma 1 3 2 as in (3 3 5)-(3 3 7), one gets

$$E_{\underline{V}_d}\{r(\hat{\underline{V}}_d, \underline{S})\hat{\underline{V}}_d \underline{Q}^* \hat{\underline{V}}_d\}$$

$$= E_{\underline{V}_d}\{(\tfrac{n}{n+2})r(\hat{\underline{V}}_d, \underline{S})\underline{V}_d \underline{Q}^* \hat{\underline{V}}_d + (\tfrac{2}{n+2})r(\hat{\underline{V}}_d, \underline{S})\underline{V}_d \operatorname{diag}(\underline{Q}^*)\hat{\underline{V}}_d$$

$$+(\tfrac{2}{n+2})\underline{V}_d[\underline{\nabla} r(\hat{\underline{V}}_d, \underline{S})]\hat{\underline{V}}_d \underline{Q}^* \hat{\underline{V}}_d\}, \qquad (3\,3\,19)$$

and

$$E_{\underline{V}_d}\{r(\hat{\underline{V}}_d, \underline{S})\underline{V}_d^{-1}\hat{\underline{V}}_d\}$$

$$= E_{\underline{V}_d}\{(\tfrac{n}{n+2})r(\hat{\underline{V}}_d, \underline{S})I_p + (\tfrac{2}{n+2})\hat{\underline{V}}_d[\underline{\nabla} r(\hat{\underline{V}}_d, \underline{S})]\} \qquad (3\,3\,20)$$

Combine (3 3.19) and (3 3 20) to get

$$E_{\underline{V}_d}\{r(\hat{\underline{V}}_d, \underline{S})\hat{\underline{V}}_d \underline{Q}^* \hat{\underline{V}}_d\}$$

$$= E_{\underline{V}_d}\{(\tfrac{n}{n+2})^2 r(\hat{\underline{V}}_d, \underline{S})\underline{V}_d \underline{Q}^* \underline{V}_d + \frac{2n}{(n+2)^2}\underline{V}_d \underline{Q}^* \underline{V}_d \hat{\underline{V}}_d[\underline{\nabla} r(\hat{\underline{V}}_d, \underline{S})]$$

$$+\frac{2n}{(n+2)^2}r(\hat{\underline{V}}_d,\underline{S})\underline{V}_d\operatorname{diag}(\underline{Q}^*)\underline{V}_d+\frac{4}{(n+2)^2}\underline{V}_d\operatorname{diag}(\underline{Q}^*)\underline{V}_d\hat{\underline{V}}_d[\underline{\nabla}r(\hat{\underline{V}}_d,\underline{S})]$$

$$+(\frac{2}{n+2})\underline{V}_d[\underline{\nabla}r(\hat{\underline{V}}_d,\underline{S})]\hat{\underline{V}}_d\underline{Q}^*\hat{\underline{V}}_d\} \qquad (3.3.21)$$

Using (3.3.11), (3.3.12) and (3.3.21), one can rewrite (3.3.18) as

$$R_{\underline{Q}^*}(\underline{\theta},\hat{\underline{\theta}}^{EB(3)}(\hat{\underline{V}}_d,\underline{X}))-R_{\underline{Q}^*}(\underline{\theta},\hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}_d,\underline{X}))$$

$$=E_{\underline{\theta},\underline{V}_d}\operatorname{tr}\{\underline{P}[\underline{\Psi}^2(\underline{\lambda})\underline{\Delta}^{-1}+2r(\hat{\underline{V}}_d,\underline{S})\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}]\underline{P}^T[(\frac{n}{n+2})^2\underline{V}_d\underline{Q}^*\underline{V}_d$$

$$+\frac{2n}{(n+2)^2}\underline{V}_d\operatorname{diag}(\underline{Q}^*)\underline{V}_d]-\underline{P}[4\underline{\Psi}^{(1)}(\underline{\lambda})+2(K-p-1)\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}]\underline{P}^T(\frac{n}{n+2})\underline{V}_d\underline{Q}^*\underline{V}_d$$

$$+2\underline{P}[r(\hat{\underline{V}}_d,\underline{S})\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}]\underline{P}^T[\frac{2n}{(n+2)^2}\underline{V}_d\underline{Q}^*\underline{V}_d\hat{\underline{V}}_d[\underline{\nabla}r(\hat{\underline{V}}_d,\underline{S})]$$

$$+\frac{4}{(n+2)^2}\underline{V}_d\operatorname{diag}(\underline{Q}^*)\underline{V}_d\hat{\underline{V}}_d[\underline{\nabla}r(\hat{\underline{V}}_d,\underline{S})]+(\frac{2}{n+2})\underline{V}_d[\underline{\nabla}r(\hat{\underline{V}}_d,\underline{S})]\hat{\underline{V}}_d\underline{Q}^*\hat{\underline{V}}_d]\} \qquad (3.3.22)$$

Suppose $\underline{Q}=\underline{Q}_d$ then $\underline{Q}^*=\operatorname{diag}(\underline{Q}^*)$ and

$$R_{\underline{Q}^*}(\underline{\theta},\hat{\underline{\theta}}^{EB(3)}(\hat{\underline{V}}_d,\underline{X}))-R_{\underline{Q}^*}(\underline{\theta},\hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}_d,\underline{X}))$$

$$\leq E_{\underline{\theta},\underline{V}_d}\operatorname{tr}\{(\frac{n}{n+2})[\underline{\Psi}^2(\underline{\lambda})\underline{\Delta}^{-1}-2\{K-p-1-r(\hat{\underline{V}}_d,\underline{S})\}\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}$$

$$-4\underline{\Psi}^{(1)}(\underline{\lambda})]\underline{P}^T\underline{V}_d^{1/2}\underline{Q}_d\underline{V}_d^{1/2}\underline{P}\} \qquad (3.3.23)$$

To show (3.3.23) < 0 for all $\underline{\theta}$ and $\underline{V}_d$, it suffices to show $\Delta_{2i}(\underline{\lambda})<0$ for each $i=1,\ldots,p$, where

$$\Delta_{2i}(\underline{\lambda}) = \Psi_i^2(\underline{\lambda})\lambda_i^{-1}-2\{(K-p-1)-r(\hat{\underline{V}}_d,\underline{S})\}\Psi_i(\underline{\lambda})\lambda_i^{-1}-2\sum_{j(\neq i)}[\Psi_i(\underline{\lambda})-\Psi_j(\underline{\lambda})]/(\lambda_i-\lambda_j)$$

$$-4[\partial\Psi_i(\underline{\lambda})/\partial\lambda_i] \qquad (3.3.24)$$

Now we prove $\hat{\underline{\theta}}^{EB(3)}(\hat{\underline{V}}_d, \underline{X})$ dominates $\hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}_d, \underline{X})$ under the L_{Q^*} loss with $Q = Q_d$ by providing solutions to (3.3.24) < 0 for each $i = 1, \ldots, p$.

Theorem 3.3.3 Let $0 < r(\hat{\underline{V}}_d, \underline{S}) < (K-2)$, $r(\hat{\underline{V}}_d, \underline{S})$ is differentiable with respect to s_{ij} for each $i, j = 1, \ldots, p$. Suppose $\Psi_i(\underline{\lambda}) = \lambda_i t(U)/(\sum_{j=1}^{p}\lambda_j)$, $i = 1, \ldots, p$, where U is a function of the λ_i's. Assume that

(i) $r(\hat{\underline{V}}_d, \underline{S})$ is non increasing in $\hat{v}_i$ for each $i = 1, \ldots, p$.

(ii) $U(\underline{\lambda})$ is non-decreasing in λ_i for each $i = 1, \ldots, p$,

(iii) $t(U)$ is non-decreasing in U,

(iv) $0 < t(U) < 2[(K-2)-r(\hat{\underline{V}}_d, \underline{S})]$.

Then $\Delta_{2i}(\underline{\lambda}) < 0$ for each $i = 1, \ldots, p$. Hence, $\hat{\underline{\theta}}^{EB(3)}(\hat{\underline{V}}_d, \underline{X})$ dominates $\hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}_d, \underline{X})$ under the L_{Q^*} loss with $Q = Q_d$.

Proof of Theorem 3.3.3 For the particular choice $\Psi_i(\underline{\lambda}) = \lambda_i t(U)/(\sum_{j=1}^{p}\lambda_j)$, one obtains from (3.3.24) that

$$\begin{aligned}\Delta_{2i}(\underline{\lambda}) = {} & \lambda_i t^2(U)/(\sum_{j=1}^{p}\lambda_j)^2 - 2[(K-p-1)-r(\hat{\underline{V}}_d, \underline{S})]t(U)/(\sum_{j=1}^{p}\lambda_j) - 2(p-1)t(U)/(\sum_{j=1}^{p}\lambda_j) \\ & -4[t(U)/(\sum_{j=1}^{p}\lambda_j) - \lambda_i t(U)/(\sum_{j=1}^{p}\lambda_j)^2 + (\lambda_i/\sum_{j=1}^{p}\lambda_j)t'(U)(\partial U/\partial\lambda_i)] \\ = {} & [t^2(U)+4t(U)]\lambda_i/(\sum_{j=1}^{p}\lambda_j)^2 - 2(K-r(\hat{\underline{V}}_d, \underline{S}))t(U)/(\sum_{j=1}^{p}\lambda_j) \\ & -4(\lambda_i/\sum_{j=1}^{p}\lambda_j)t'(U)(\partial U/\partial\lambda_i)\end{aligned}$$

$$\le \; [t^2(U) - 2\{(K-2) - r(\hat{\underline{V}}_d, \underline{S})\}t(U)]/(\sum_{j=1}^{p}\lambda_j) < 0 \qquad (3.3.25)$$

under the assumptions (i)-(iii). The conclusion of this theorem followed.

If we concentrate on the improvement over $(\underline{I}_p - a\hat{\underline{V}}_d\underline{S}^{-1})\underline{X}$, that is, $r(\hat{\underline{V}}_d, \underline{S}) = a$ is a constant function, then $(\underline{I}_p - a\hat{\underline{V}}_d\underline{S}^{-1} - [t(U)/tr(\underline{S})]\hat{\underline{V}}_d)\underline{X}$ dominates $(\underline{I}_p - a\hat{\underline{V}}_d\underline{S}^{-1})\underline{X}$ under the L_{Q^*} loss with any p.d. $\underline{Q}$ is shown in the following theorem.

Theorem 3.3.4 Let $0 < a < r_0(K-2)$ where $r_0 = (n+2)/[n+2(\max q_{ii})/(\min \eta_i)]$. Suppose $\Psi_i(\underline{\lambda}) = \lambda_i t(U)/(\sum_{j=1}^{p}\lambda_j)$, $i = 1, \ldots, p$, where U is a function of the λ_i's. $(\underline{I}_p - a\hat{\underline{V}}_d\underline{S}^{-1} - [t(U)/tr(\underline{S})]\hat{\underline{V}}_d)\underline{X}$ dominates $(\underline{I}_p - a\hat{\underline{V}}_d\underline{S}^{-1})\underline{X}$ under the L_{Q^*} loss with any p.d. $\underline{Q}$ if

(i) $U(\underline{\lambda})$ is non-decreasing in λ_i for each $i = 1, \ldots, p$;

(ii) $t(U)$ is non-decreasing in U,

(iii) $0 < t(U) < 2[r_0(K-2) - a]$.

Proof of Theorem 3.3.4 In view of (3.3.22),

$$R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(3)}(\hat{\underline{V}}_d, \underline{X})) - R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}_d, \underline{X}))$$

$$\le E_{\underline{\theta}, \underline{V}_d} tr[\underline{P}\{\underline{\Psi}^2(\underline{\lambda})\underline{\Delta}^{-1} + 2a\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}\}\underline{P}^T[(\tfrac{n}{n+2})^2$$

$$+ \frac{2n}{(n+2)^2}((\max q_{ii})/(\min \eta_i))]\underline{V}_d^{1/2}\underline{Q}\underline{V}_d^{1/2}$$

$$-\underline{P}\{4\underline{\Psi}^{(1)}(\underline{\lambda}) + 2(K-p-1)\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}\}\underline{P}^T(\tfrac{n}{n+2})\underline{V}_d^{1/2}\underline{Q}\underline{V}_d^{1/2}\}$$

$$= E_{\underline{\theta}, \underline{V}_d} tr\{\frac{n}{(n+2)} r_0^{-1}[\underline{\Psi}^2(\underline{\lambda})\underline{\Delta}^{-1} - 2[r_0(K-p-1) - a]\underline{\Psi}(\underline{\lambda})\underline{\Delta}^{-1}$$

$$-4r_0\underline{\Psi}^{(1)}(\underline{\lambda})[\underline{P}^T\underline{V}_d Q^* \underline{V}_d \underline{P}\} \tag{3.3.26}$$

To show (3.3.26) < 0 for all $\underline{\theta}$ and $\underline{V}_d$ is suffices to show $\Delta_{2'i}(\underline{\lambda}) < 0$ for each $i = 1, \ldots, p$, where

$$\Delta_{2'i}(\underline{\lambda}) = \Psi_i^2(\underline{\lambda})\lambda_i^{-1} - 2[r_0(K-p-1)-a]\Psi_i(\underline{\lambda})\lambda_i^{-1} - 2r_0 \sum_{j(\neq i)} [\Psi_i(\underline{\lambda}) - \Psi_j(\underline{\lambda})]/(\lambda_i - \lambda_j)$$
$$-4r_0[\partial\Psi_i(\underline{\lambda})/\partial\lambda_i] \tag{3.3.27}$$

With the choice of $\Psi_i(\underline{\lambda}) = \lambda_i t(U)/(\sum_{j=1}^{p}\lambda_j)$, one obtains from (3.3.27) that

$$\Delta_{2'i}(\underline{\lambda}) = \lambda_i t^2(U)/(\sum_{j=1}^{p}\lambda_j)^2 - 2[r_0(K-p-1)-a]t(U)/(\sum_{j=1}^{p}\lambda_j) - 2r_0(p-1)t(U)/(\sum_{j=1}^{p}\lambda_j)$$
$$-4r_0[t(U)/(\sum_{j=1}^{p}\lambda_j) - \lambda_i t(U)/(\sum_{j=1}^{p}\lambda_j)^2 + (\lambda_i/\sum_{j=1}^{p}\lambda_j)t'(U)(\partial U/\partial\lambda_i)]$$
$$= [t^2(U) + 4r_0 t(U)]\lambda_i/(\sum_{j=1}^{p}\lambda_j)^2 - 2(r_0K-a)t(U)/(\sum_{j=1}^{p}\lambda_j)$$
$$-4r_0(\lambda_i/\sum_{j=1}^{p}\lambda_j)t'(U)(\partial U/\partial\lambda_i)$$
$$\leq [t^2(U) - 2[r_0(K-2)-a]t(U)]/(\sum_{j=1}^{p}\lambda_j) < 0, \tag{3.3.28}$$

under the assumptions (i)-(iii) and the proof is complete.

Remark 3.3.5 For the particular choice of $a = r_0(K-p-1)$, $(\underline{I}_p - a\hat{\underline{V}}_d\underline{S}^{-1} - [t(U)/tr(\underline{S})]\hat{\underline{V}}_d)\underline{X}$ dominates $(\underline{I}_p - a\hat{\underline{V}}_d\underline{S}^{-1})\underline{X}$ under the L_{Q^*} loss with any p.d. Q whenever $0 < t(U) < 2r_0(p-1)$. Furthermore, if $t(U) = b$ is also constant, it can be shown that the optimal choice is $b = r_0(p-1)$. Notice that if Q is a diagonal matrix, then Theorem 3.3.4 can be slightly extended by letting $r_0 = 1$.

Remark 3.3.6 Consider the L_Q loss instead of L_{Q^*} loss in (3.3.18) (3.3.22). The conditions in Theorem 3.3.3 and Theorem 3.3.4 still support $R_Q(\underline{\theta}, \widehat{\underline{\theta}}^{EB(3)}(\widehat{\underline{V}}_\phi, \underline{X})) < R_Q(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)}(\widehat{\underline{V}}_\phi, \underline{X}))$.

3.4 EB Minimax Estimators Under the L_{Q^*} loss and $\underline{V}$ Is Completely Unknown

In this section, EB estimators of $\underline{\theta}$ are discussed under the L_{Q^*} loss, and the covariance matrix $\underline{V}$ of $\underline{X}_i$'s is complete unknown and needs to be estimated from the supplementary data. We assume $\widehat{\underline{V}} \sim W(p, n, (n+p+1)^{-1}\underline{V})$ is available as in (3.1.10). In a balanced one-way MANOVA model introduced in (2.1.2), suppose $\underline{V}_0$ is entirely unknown and $n = K(m-1)$. Then

$$\widehat{\underline{V}} = m^{-1}\sum_{i=1}^{K}\sum_{j=1}^{m}(\underline{X}_{ij}-\underline{X}_i)(\underline{X}_{ij}-\underline{X}_i)^T/(n+p+1)$$

is the best multiple estimator of $\underline{V} = m^{-1}\underline{V}_0$. Note that conditional on $\underline{\theta}$, $\underline{X}_i$'s and $\widehat{\underline{V}}$ are independent, and the distribution of $\widehat{\underline{V}}$ does not depend on $\underline{\theta}$. Also marginally, $(\underline{S}, \widehat{\underline{V}})$ is the complete sufficient statistic for $(\underline{\Sigma}, \underline{V})$.

The two classes of EB estimators $\widehat{\underline{\theta}}^{EB(1)}(\widehat{\underline{V}}, \underline{X})$ and $\widehat{\underline{\theta}}^{EB(2)}(\widehat{\underline{V}}, \underline{X})$ defined in (3.1.11) and (3.1.12) respectively are going to be used in this section. Estimators $\widehat{\underline{\theta}}^{EB(1)}$ of the form (3.1.11) are direct extensions of Baranchik's (1970) type of estimators for the matrix normal means and complete unknown covariance matrix, while estimators $\widehat{\underline{\theta}}^{EB(2)}$ defined in (3.1.12) generalize those in Efron and Morris (1972b, 1976b), Stein (1974), Dey, Ghosh and Srinivasan (1985) and Zheng (1986) for complete unknown covariance. Both classes of estimators are investigated under the general quadratic loss L_{Q^*} given in (3.1.1) and the generalized version of the James-Stein estimator namely $(\underline{I}_p - a\widehat{\underline{V}}\underline{S}^{-1})\underline{X}$ is included as a member in both classes where a is some constant.

Recall an EB estimator $\widehat{\underline{\theta}}^{EB}$ has risk $R_{Q^*}(\underline{\theta}, \widehat{\underline{\theta}}^{EB})$ expressed in (3.1.3). Based on the form of $\widehat{\underline{\theta}}^{EB(1)}$, one way of expressing the risk $R_{Q^*}(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)})$ is given in the following corollary.

Corollary 3.4.1 Under the L_{Q^*} loss, $\widehat{\underline{\theta}}^{EB(1)}(\widehat{\underline{V}}, \underline{X}) = (I_p - r(\widehat{\underline{V}}, \underline{S})\widehat{\underline{V}}\underline{S}^{-1})\underline{X}$ has risk

$$R_{Q^*}(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)})$$

$$= R_{Q^*}(\underline{\theta}, \underline{X}) + n(n+p+1)^{-2}E_{\underline{\theta},\underline{V}}\mathrm{tr}\{(n+1)r^2(\widehat{\underline{V}}, \underline{S})\underline{S}^{-1}\underline{V}Q^*\underline{V} + r^2(\widehat{\underline{V}}, \underline{S})\mathrm{tr}(Q^*\underline{V})\underline{S}^{-1}\underline{V}$$

$$-2(n+p+1)(K-p-1)r(\widehat{\underline{V}}, \underline{S})\underline{S}^{-1}\underline{V}Q^*\underline{V}\} + 2(n+p+1)^{-2}E_{\underline{\theta},\underline{V}}\mathrm{tr}\{[D^{\widehat{\underline{V}}}r^2(\widehat{\underline{V}}, \underline{S})]$$

$$\cdot\{(n+p+1)\widehat{\underline{V}}\underline{S}^{-1}\widehat{\underline{V}}Q^*\underline{V} + \widehat{\underline{V}}\mathrm{tr}(Q^*\underline{V})\underline{S}^{-1}\underline{V} + (n+1)\widehat{\underline{V}}\underline{S}^{-1}\underline{V}Q^*\underline{V}\}$$

$$-2(n+p+1)(K-p-1)[D^{\widehat{\underline{V}}}r(\widehat{\underline{V}}, \underline{S})]\widehat{\underline{V}}\underline{S}^{-1}\underline{V}Q^*\underline{V}]$$

$$-4E_{\underline{\theta},\underline{V}}\mathrm{tr}\{[D^{\underline{S}}r(\widehat{\underline{V}}, \underline{S})]\widehat{\underline{V}}Q^*\underline{V}\} \tag{3.4.1}$$

Proof of Corollary 3.4.1 By expressing $\underline{r}(\widehat{\underline{V}}, \underline{S}) = r(\widehat{\underline{V}}, \underline{S})I_p$ in (3.1.3), one gets

$$R_{Q^*}(\underline{\theta}, \widehat{\underline{\theta}}^{EB(1)})$$

$$= R_{Q^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta},\underline{V}}\mathrm{tr}\{r^2(\widehat{\underline{V}}, \underline{S})\underline{S}^{-1}\widehat{\underline{V}}Q^*\widehat{\underline{V}} - 4[D^{\underline{S}}r(\widehat{\underline{V}}, \underline{S})]\widehat{\underline{V}}Q^*\underline{V}$$

$$-2(K-p-1)r(\widehat{\underline{V}}, \underline{S})\underline{S}^{-1}\widehat{\underline{V}}Q^*\underline{V}\} \tag{3.4.2}$$

Recall $\widehat{\underline{V}} \sim W(p, n, (n+p+1)^{-1}\underline{V})$ is independent of $\underline{X}$, and define $D^{\widehat{\underline{V}}} = ((d_{ij}^{\widehat{V}}))$ as in Lemma 1.3.3. Applying the Wishart identity in Lemma (1.3.3), we have

$$E_{\underline{V}}\{r(\widehat{\underline{V}}, \underline{S})\widehat{\underline{V}}\}$$

$$= E_{\underline{V}}\{(n+p+1)^{-1}\underline{V}[(n+p+1)^{-1}\underline{V})^{-1}r(\widehat{\underline{V}}, \underline{S})\widehat{\underline{V}}]$$

$$= (n+p+1)^{-1}\underline{V}\,E_{\underline{V}}\{2\underline{D}^{\hat{\underline{V}}}[r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}]+(n-p-1)r(\hat{\underline{V}}, \underline{S})\underline{I}_p\} \tag{3.4.3}$$

By performing the formal multiplication and differentiating coordinatewise (see Lemma 1, Haff(1981)), we get

$$\underline{D}^{\hat{\underline{V}}}[r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}]$$

$$= [\underline{D}^{\hat{\underline{V}}}r(\hat{\underline{V}}, \underline{S})]\hat{\underline{V}}+[(r(\hat{\underline{V}}, \underline{S})\underline{I}_p)^T(\underline{D}^{\hat{\underline{V}}})^T]^T\hat{\underline{V}}$$

$$= [\underline{D}^{\hat{\underline{V}}}r(\hat{\underline{V}}, \underline{S})]\hat{\underline{V}}+r(\hat{\underline{V}}, \underline{S})[\underline{D}^{\hat{\underline{V}}}\hat{\underline{V}}], \tag{3.4.4}$$

also

$$[\underline{D}^{\hat{\underline{V}}}\hat{\underline{V}}] = ((\sum_k \tfrac{1}{2}(1+\delta_{ik})(\partial\hat{v}_{kj}/\partial\hat{v}_{ik})))$$

$$= ((\sum_k \tfrac{1}{2}(1+\delta_{ik})\delta_{ij}))$$

$$= ((\tfrac{1}{2}(p+1)\delta_{ij}))$$

$$= \tfrac{1}{2}(p+1)\underline{I}_p \tag{3.4.5}$$

Combine (3.4.3)-(3.4.5) to obtain

$$E_{\underline{V}}\{r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\} = (n+p+1)^{-1}\underline{V}\,E_{\underline{V}}\{2[\underline{D}^{\hat{\underline{V}}}r(\hat{\underline{V}}, \underline{S})]\hat{\underline{V}}+n\,r(\hat{\underline{V}}, \underline{S})\underline{I}_p\} \tag{3.4.6}$$

Therefore,

$$E_{\underline{\theta},\underline{V}}\mathrm{tr}\{r(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^*\underline{V}\}$$

$$= E_{\underline{\theta},\underline{V}}\mathrm{tr}\{\underline{S}^{-1}r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}[(n+p+1)^{-1}\underline{V}]^{-1}[(n+p+1)^{-1}\underline{V}]\underline{Q}^*\underline{V}\}$$

$$= (n+p+1)^{-1}E_{\underline{\theta},\underline{V}}\mathrm{tr}\{\underline{V}\,\underline{Q}^*\underline{V}\,\underline{S}^{-1}r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}[(n+p+1)^{-1}\underline{V}]^{-1}\}$$

$$= (n+p+1)^{-1} E_{\underline{\theta}\,\underline{V}} \mathrm{tr}\{\underline{V}\,\underline{Q}^{*}\underline{V}\,\underline{S}^{-1} r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}[(n+p+1)^{-1}\underline{V}]^{-1}\}^{T}$$

$$= (n+p+1)^{-1} E_{\underline{\theta}\,\underline{V}} \mathrm{tr}\{[(n+p+1)^{-1}\underline{V}]^{-1} r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\,\underline{S}^{-1}\underline{V}\,\underline{Q}^{*}\underline{V}\}$$

$$= (n+p+1)^{-1} E_{\underline{\theta}\,\underline{V}} \mathrm{tr}\{2[\underline{D}^{\hat{\underline{V}}} r(\hat{\underline{V}}, \underline{S})]\hat{\underline{V}}\,\underline{S}^{-1}\underline{V}\,\underline{Q}^{*}\underline{V} + n\, r(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\underline{V}\,\underline{Q}^{*}\underline{V}\} \tag{3.4.7}$$

Next, we consider the expression

$$E_{\underline{\theta}\,\underline{V}} \mathrm{tr}\{r^{2}(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^{*}\hat{\underline{V}}\}$$

$$= E_{\underline{\theta}\,\underline{V}} \mathrm{tr}\{r^{2}(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^{*}\}$$

$$= (n+p+1)^{-1} E_{\underline{\theta}\,\underline{V}} \mathrm{tr}\{[(n+p+1)^{-1}\underline{V}]^{-1} r^{2}(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^{*}\underline{V}\} \tag{3.4.8}$$

Lemma 1.3.3 is applied again for $\hat{\underline{V}} \sim W(p, n, (n+p+1)^{-1}\underline{V})$ to get

$$E_{\underline{V}}[[(n+p+1)^{-1}\underline{V}]^{-1} r^{2}(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}}]$$

$$= E_{\underline{V}}\{2\underline{D}^{\hat{\underline{V}}}[r^{2}(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}}] + (n-p-1) r^{2}(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\}, \tag{3.4.9}$$

where

$$E_{\underline{V}}\{\underline{D}^{\hat{\underline{V}}}[r^{2}(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}}]\}$$

$$= E_{\underline{V}}\{[\underline{D}^{\hat{\underline{V}}} r^{2}(\hat{\underline{V}}, \underline{S})]\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}} + [(r^{2}(\hat{\underline{V}}, \underline{S})\underline{I}_{p})^{T}(\underline{D}^{\hat{\underline{V}}})^{T}]^{T}\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}}\}$$

$$= E_{\underline{V}}\{[\underline{D}^{\hat{\underline{V}}} r^{2}(\hat{\underline{V}}, \underline{S})]\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}} + r^{2}(\hat{\underline{V}}, \underline{S})\underline{D}^{\hat{\underline{V}}}(\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}})\}, \tag{3.4.10}$$

and

$$\underline{D}^{\hat{\underline{V}}}(\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}})$$

$$= [\underline{D}^{\hat{\underline{V}}}\hat{\underline{V}}]\underline{S}^{-1}\hat{\underline{V}} + [(\hat{\underline{V}}^{T}(\underline{D}^{\hat{\underline{V}}})^{T})^{T}\underline{S}^{-1}]\hat{\underline{V}} + [(\underline{S}^{-1})^{T}\hat{\underline{V}}^{T}(\underline{D}^{\hat{\underline{V}}})^{T}]^{T}\hat{\underline{V}}$$

$$= [(\tfrac{p+1}{2})\underline{I}_{p}]\underline{S}^{-1}\hat{\underline{V}} + \underline{0} + [\tfrac{1}{2}\mathrm{tr}(\underline{S}^{-1}\hat{\underline{V}})\underline{I}_{p} + \tfrac{1}{2}\underline{S}^{-1}\hat{\underline{V}}]$$

$$= (\tfrac{n+2}{2})\underline{S}^{-1}\hat{\underline{V}} + \tfrac{1}{2}\mathrm{tr}(\underline{S}^{-1}\hat{\underline{V}}) \qquad (3.4.11)$$

Both (3.4.10) and (3.4.11) use the identities in Lemma 1 of Haff (1981). We have $[\underline{D}^{\hat{V}}\hat{\underline{V}}] = (p+1)/2\underline{I}_p$ from (3.4.5) and $[(\underline{S}^{-1})^T\hat{\underline{V}}^T(\underline{D}^{\hat{V}})^T]^T\hat{\underline{V}} = 2^{-1}[\mathrm{tr}(\underline{S}^{-1}\hat{\underline{V}})\underline{I}_p + \underline{S}^{-1}\hat{\underline{V}}]$ from Lemma 3 of Haff (1982). From (3.4.8)-(3.4.11), one gets

$$E_{\underline{\theta}}\underline{Y}\mathrm{tr}\{r^2(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^*\hat{\underline{V}}\}$$

$$= (n+p+1)^{-1}E_{\underline{\theta}}\underline{Y}\mathrm{tr}\{2[\underline{D}^{\hat{V}}r^2(\hat{\underline{V}}, \underline{S})]\hat{\underline{V}}\,\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^*\underline{Y} + r^2(\hat{\underline{V}}, \underline{S})\mathrm{tr}(\underline{S}^{-1}\hat{\underline{V}})\underline{Q}^*\underline{Y}$$

$$+ (n+1)r^2(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^*\underline{Y}\} \qquad (3.4.12)$$

Now the last two terms in (3.4.12) can be written as

$$E_{\underline{\theta}}\underline{Y}\mathrm{tr}\{r^2(\hat{\underline{V}}, \underline{S})\mathrm{tr}(\underline{S}^{-1}\hat{\underline{V}})\underline{Q}^*\underline{Y}\}$$

$$= E_{\underline{\theta}}\underline{Y}\mathrm{tr}(\underline{S}^{-1}\hat{\underline{V}})\mathrm{tr}\{r^2(\hat{\underline{V}}, \underline{S})\underline{Q}^*\underline{Y}\}$$

$$= E_{\underline{\theta}}\underline{Y}\mathrm{tr}\{r^2(\hat{\underline{V}}, \underline{S})\mathrm{tr}(\underline{Q}^*\underline{Y})\underline{S}^{-1}\hat{\underline{V}}\}$$

$$= E_{\underline{\theta}}\underline{Y}\mathrm{tr}\{r^2(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\mathrm{tr}(\underline{Q}^*\underline{Y})\underline{S}^{-1}\}, \qquad (3.4.13)$$

and

$$E_{\underline{\theta}}\underline{Y}\mathrm{tr}\{r^2(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^*\underline{Y}\}$$

$$= E_{\underline{\theta}}\underline{Y}\mathrm{tr}\{r^2(\hat{\underline{V}}, \underline{S})\underline{Q}^*\underline{Y}\,\underline{S}^{-1}\hat{\underline{V}}\}$$

$$= E_{\underline{\theta}}\underline{Y}\mathrm{tr}\{r^2(\hat{\underline{V}}, \underline{S})\underline{Q}^*\underline{Y}\,\underline{S}^{-1}\hat{\underline{V}}\}^T$$

$$= E_{\underline{\theta}}\underline{Y}\mathrm{tr}\{r^2(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\,\underline{S}^{-1}\underline{Y}\underline{Q}^*\}. \qquad (3.4.14)$$

Also

$$E_{\underline{V}}\{r^2(\hat{\underline{V}},\underline{S})\hat{\underline{V}}\} = (n+p+1)^{-1}\underline{V}\,E_{\underline{V}}\{2[\underline{D}^{\hat{\underline{V}}}r^2(\hat{\underline{V}},\underline{S})]\hat{\underline{V}}+n\,r^2(\hat{\underline{V}},\underline{S})\}, \qquad (3.4.15)$$

as in (3.4.6) by replacing $r(\hat{\underline{V}},\underline{S})$ with $r^2(\hat{\underline{V}},\underline{S})$.

Combing (3.4.12)-(3.4.15), one has

$$E_{\underline{\theta},\underline{V}}\mathrm{tr}\{r^2(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^*\hat{\underline{V}}\}$$

$$= (n+p+1)^{-2}E_{\underline{\theta},\underline{V}}\mathrm{tr}\{2(n+p+1)[\underline{D}^{\hat{\underline{V}}}r^2(\hat{\underline{V}},\underline{S})]\hat{\underline{V}}\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^*\underline{V}$$

$$+2[\underline{D}^{\hat{\underline{V}}}r^2(\hat{\underline{V}},\underline{S})]\hat{\underline{V}}\mathrm{tr}(\underline{Q}^*\underline{V})\underline{S}^{-1}\underline{V}+n\,r^2(\hat{\underline{V}},\underline{S})\mathrm{tr}(\underline{Q}^*\underline{V})\underline{S}^{-1}\underline{V}$$

$$+2(n+1)[\underline{D}^{\hat{\underline{V}}}r^2(\hat{\underline{V}},\underline{S})]\hat{\underline{V}}\underline{S}^{-1}\underline{V}\underline{Q}^*\underline{V}+n(n+1)r^2(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\underline{V}\underline{Q}^*\underline{V}\} \qquad (3.4.16)$$

Substituting (3.4.7) and (3.4.16) into (3.4.2) to get (3.4.1), the proof is complete.

Consider the particular choice $r(\hat{\underline{V}},\underline{S}) = r(F)$ where $F = \mathrm{tr}(\underline{S}^{-1}\hat{\underline{V}})$. Then $[\underline{D}^{\underline{S}}r(\hat{\underline{V}},\underline{S})] = r'(F)[\underline{D}^{\underline{S}}F] = r'(F)\hat{\underline{V}}^{-1}$, $[\underline{D}^{\hat{\underline{V}}}r(\hat{\underline{V}},\underline{S})] = r'(F)[\underline{D}^{\hat{\underline{V}}}F]$ and $[\underline{D}^{\hat{\underline{V}}}r^2(\hat{\underline{V}},\underline{S})] = 2r(F)\cdot r'(F)[\underline{D}^{\hat{\underline{V}}}F]$ where $[\underline{D}^{\hat{\underline{V}}}F] = [\underline{D}^{\hat{\underline{V}}}\mathrm{tr}(\underline{S}^{-1}\hat{\underline{V}})] = -\hat{\underline{V}}^{-1}\underline{S}\hat{\underline{V}}^{-1}$, from Lemma 6 of Haff (1982). We have the following theorem.

Theorem 3.4.1 Let $r(\hat{\underline{V}},\underline{S}) = r(F)$, where $F = \mathrm{tr}(\underline{S}^{-1}\hat{\underline{V}})$. Then estimators of the form $\hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}},\underline{X}) = (\underline{I}_p - r(F)\hat{\underline{V}}\underline{S}^{-1})\underline{X}$ dominate $\underline{X}$ under the $L_{\underline{Q}^*}$ loss if

(i) $0 < r(F) < 2r_0(Kp-1)$

(ii) $r(F)$ is non-decreasing in F;

(iii) $[\underline{D}^{\hat{\underline{V}}}r']$ is n.n.d.,

where $r_0 = (n+p+1)/[n+\mathrm{tr}(\underline{Q})/(\min \eta_i)+1]$.

Proof of Theorem 3.4.1 By applying $[\underline{D}^{\hat{\underline{V}}}r^2(F)] = -2r(F)\cdot r'(F)\hat{\underline{V}}^{-1}\underline{S}\hat{\underline{V}}^{-1}$, we have

$$\operatorname{tr}\{[D^{\hat{\underline{V}}} r^2(F)][(n+p+1)\hat{\underline{V}}\underline{S}^{-1}\hat{\underline{V}}Q^*\underline{V}+\hat{\underline{V}}\operatorname{tr}(Q^*\underline{V})\underline{S}^{-1}\underline{V}+(n+1)\hat{\underline{V}}\underline{S}^{-1}\underline{V}Q^*\underline{V}]\}$$

$$= -2r(F)r'(F)\operatorname{tr}\{(n+p+1)Q+\operatorname{tr}(Q)\hat{\underline{V}}^{-1/2}\underline{V}\hat{\underline{V}}^{-1/2}+(n+1)\hat{\underline{V}}^{-1/2}V^{-1/2}QV^{-1/2}\hat{\underline{V}}^{-1/2}]$$

$$\leq 0, \tag{3.4.17}$$

The inequality follows from the assumption (ii) and all $\hat{\underline{V}}$, $\underline{V}$ and Q are p d. Also,

$$\operatorname{tr}\{[D^{\underline{S}} r(F)]\hat{\underline{V}}Q^*\underline{V}\} = \operatorname{tr}\{r'(F)\hat{\underline{V}}^{-1}\hat{\underline{V}}Q^*\underline{V}\} = r'(F)\operatorname{tr}(Q) \geq 0, \tag{3.4.18}$$

since $[D^{\underline{S}} r(F)] = r'(F)[D^{\underline{S}} F] = r'(F)\hat{\underline{V}}^{-1}$, $r(F)$ is non-decreasing in F and Q is p d..

Using $[D^{\hat{\underline{V}}} r(F)] = -r'(F)\hat{\underline{V}}^{-1}\underline{S}\,\hat{\underline{V}}^{-1}$ to get $\operatorname{tr}\{[D^{\hat{\underline{V}}} r(F)]\hat{\underline{V}}\underline{S}^{-1}\underline{V}Q^*\underline{V}\} = \operatorname{tr}\{-r'(F)\hat{\underline{V}}^{-1}\underline{V}Q^*\underline{V}\}$ From the Wishart identity

$$E_{\underline{V}}\{\underline{V}^{-1}r'(F)\} = E_{\underline{V}}\{2[D^{\hat{\underline{V}}} r'(F)]+(n-p-1)\hat{\underline{V}}^{-1}r'(F)\}, \tag{3.4.19}$$

one can show

$$E_{\underline{V}}\operatorname{tr}\{r'(F)\hat{\underline{V}}^{-1}\underline{V}Q^*\underline{V}\}$$

$$= (n-p-1)^{-1}E_{\underline{V}}\operatorname{tr}\{r'(F)Q-2Q^{1/2}\underline{V}^{1/2}[D^{\hat{\underline{V}}} r'(F)]\underline{V}^{1/2}Q^{1/2}\} \tag{3.4.20}$$

Hence form (3.4.18) and (3.4.20), one has

$$(n+p+1)^{-1}(K-p-1)\operatorname{tr}\{[D^{\hat{\underline{V}}} r(\hat{\underline{V}}, \underline{S})]\hat{\underline{V}}\underline{S}^{-1}\underline{V}Q^*\underline{V}\}+\operatorname{tr}\{[D^{\underline{S}} r(\hat{\underline{V}}, \underline{S})]\hat{\underline{V}}Q^*\underline{V}\}$$

$$= \{1-((n+p+1)(n-p-1))^{-1}(K-p-1)\}r'(F)\operatorname{tr}(Q)$$

$$+2((n+p+1)(n-p-1))^{-1}(K-p-1)\operatorname{tr}\{Q^{1/2}\underline{V}^{1/2}[D^{\hat{\underline{V}}} r'(F)]\underline{V}^{1/2}Q^{1/2}\} \geq 0, \tag{3.4.21}$$

since $[D^{\hat{\underline{V}}} r'(F)]$ is assumed n.n.d., and n is usually larger than K and r(F) is non-decreasing in F From (3.4.1), (3.4.17) and (3.4.21), it follows that

$$R_{Q^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)})-R_{Q^*}(\theta, \underline{X})$$

$$\leq \; n(n+p+1)^{-2}E_{\underline{\theta},\underline{V}}\mathrm{tr}\{(n+1)r^2(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\underline{V}\,\underline{Q}^*\underline{V}+r^2(\hat{\underline{V}},\underline{S})\mathrm{tr}(\underline{Q}^*\underline{V})\underline{S}^{-1}\underline{V}$$

$$-2(n+p+1)(K-p-1)r(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\underline{V}\,\underline{Q}^*\underline{V}) \tag{3.4.22}$$

Recall that $(\min \eta_i)$ is the smallest eigen value of $\underline{Q}$ and $\underline{S}^{-1/2}\underline{V}^{1/2}\underline{Q}\,\underline{V}^{1/2}\underline{S}^{-1/2}$ is p.d. Using $\mathrm{tr}(\underline{Q}^*\underline{V})\underline{I}_p = \mathrm{tr}(\underline{Q})\underline{I}_p \leq (\mathrm{tr}(\underline{Q})/(\min \eta_i))\underline{Q}$, one gets $\mathrm{tr}(\mathrm{tr}(\underline{Q}^*\underline{V})\,\underline{S}^{-1}\underline{V}) = \mathrm{tr}(\underline{S}^{-1/2}\underline{V}^{1/2}\mathrm{tr}(\underline{Q})\underline{V}^{1/2}\underline{S}^{-1/2}) \leq \mathrm{tr}(\underline{S}^{-1/2}\underline{V}^{1/2}[\mathrm{tr}(\underline{Q})/(\min \eta_i)]\underline{Q}\,\underline{V}^{1/2}\underline{S}^{-1/2})$. Thus we have the following inequality

$$R_{\underline{Q}^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}) - R_{\underline{Q}^*}(\underline{\theta}, \underline{X})$$

$$\leq \; n(n+p+1)^{-2}E_{\underline{\theta},\underline{V}}\mathrm{tr}\{[(n+\mathrm{tr}(\underline{Q})/(\min \eta_i)+1)r^2(\hat{\underline{V}},\underline{S})-2(n+p+1)(K-p-1)r(\hat{\underline{V}},\underline{S})]$$

$$\underline{S}^{-1/2}\underline{V}^{1/2}\underline{Q}\,\underline{V}^{1/2}\underline{S}^{-1/2}\} \tag{3.4.23}$$

Then the result is a consequence of assumption (i). The proof is complete.

Remark 3.4.1 Suppose $r(\hat{\underline{V}},\underline{S}) = a$ is constant. In view of from (3.4.23),

$$R_{\underline{Q}^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(1)}) - R_{\underline{Q}^*}(\underline{\theta}, \underline{X})$$

$$= \; n(n+p+1)^{-1}r_0^{-1}E_{\underline{\theta},\underline{V}}\mathrm{tr}\{(a^2-2r_0(K-p-1)a)\underline{S}^{-1/2}\underline{V}^{1/2}\underline{Q}\,\underline{V}^{1/2}\underline{S}^{-1/2}\} \tag{3.4.24}$$

It can be shown that the best value of a is $r_0(K-p-1)$ and the corresponding estimator is the matrix form of James-Stein estimator under the $L_{\underline{Q}^*}$ loss and $\underline{V}$ complete unknown, that is, $(\underline{I}_p - r_0(K-p-1)\hat{\underline{V}}\,\underline{S}^{-1})\underline{X}$.

Next, we consider the family of estimators of the form $\hat{\underline{\theta}}^{EB(2)}$ defined in (3.1.12). The following corollary provides the risk for $\hat{\underline{\theta}}^{EB(2)}$ under the $L_{\underline{Q}^*}$ loss in (3.1.1). The risk is denoted by $R_{\underline{Q}^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(2)})$.

Corollary 3.4.2 Under the $L_{\underline{Q}^*}$ loss, $\underline{\hat{\theta}}^{EB(2)}(\underline{\hat{V}}, \underline{X}) = (\underline{I}_p - \underline{\hat{V}}\,\underline{S}^{-1}\underline{g}(\underline{S}))\underline{X}$ has risk

$$
\begin{aligned}
& R_{\underline{Q}^*}(\underline{\theta}, \underline{\hat{\theta}}^{EB(2)}) \\
= \; & R_{\underline{Q}^*}(\underline{\theta}, \underline{X}) + n(n+p+1)^{-2} E_{\underline{\theta},\underline{V}} \operatorname{tr}[[\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1} - 4(n+p+1)\underline{D}^{\underline{S}}[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}] \\
& - 2(n+p+1)(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}]\underline{V}^{1/2}\underline{Q}\,\underline{V}^{1/2} + \underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}\operatorname{tr}(\underline{Q})\underline{V}] \qquad (3.4.25)
\end{aligned}
$$

Proof of Corollary 3.4.2 With the change of $\underline{r}(\underline{\hat{V}}, \underline{S})$ with $\underline{g}(\underline{S})$ in (3.1.3), we obtain

$$
\begin{aligned}
& R_{\underline{Q}^*}(\underline{\theta}, \underline{\hat{\theta}}^{EB(2)}) \\
= \; & R_{\underline{Q}^*}(\underline{\theta}, \underline{X}) + E_{\underline{\theta},\underline{V}} \operatorname{tr}\{\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}\underline{\hat{V}}\,\underline{Q}^*\underline{\hat{V}} - 4\underline{D}^{\underline{S}}[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}]\underline{\hat{V}}\,\underline{Q}^*\underline{V} \\
& - 2(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}\underline{\hat{V}}\,\underline{Q}^*\underline{V}\} \qquad (3.4.26)
\end{aligned}
$$

Since $\underline{\hat{V}} \sim W(p, n, (n+p+1)^{-1}\underline{V})$, by applying Lemma 1.3.3 twice to get

$$
E_{\underline{V}}[\underline{\hat{V}}\,\underline{Q}^*\underline{\hat{V}}] = n(n+p+1)^{-2}[(n+1)\underline{V}\,\underline{Q}^*\underline{V} + \operatorname{tr}(\underline{V}\,\underline{Q}^*)\underline{V}] \qquad (3.4.27)
$$

Also $E_{\underline{V}}[\underline{\hat{V}}] = n(n+p+1)^{-1}\underline{V}$.

In view of (3.4.26), one has

$$
\begin{aligned}
& R_{\underline{Q}^*}(\underline{\theta}, \underline{\hat{\theta}}^{EB(2)}) \\
= \; & R_{\underline{Q}^*}(\underline{\theta}, \underline{X}) + n(n+p+1)^{-2} E_{\underline{\theta},\underline{V}} \operatorname{tr}\{\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}[(n+1)\underline{V}\,\underline{Q}^*\underline{V} + \operatorname{tr}(\underline{V}\,\underline{Q}^*)\underline{V}] \\
& - 4(n+p+1)\underline{D}^{\underline{S}}[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}]\underline{V}\,\underline{Q}^*\underline{V} - 2(n+p+1)(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}\underline{V}\,\underline{Q}^*\underline{V}\} \qquad (3.4.28)
\end{aligned}
$$

Equation (3.4.25) is a direct consequence of (3.4.27) and $\underline{Q}^* = \underline{V}^{-1/2}\underline{Q}\,\underline{V}^{-1/2}$. The proof is complete.

Note $\mathrm{tr}(\underline{Q})\underline{V} \le (\mathrm{tr}(\underline{Q})/(\min \eta_i))\underline{V}^{1/2}\underline{Q}\,\underline{V}^{1/2}$. The risk difference of $\hat{\underline{\theta}}^{EB(2)}$ and $\underline{X}$ is

$$R_{\underline{Q}^*}(\underline{\theta}, \hat{\underline{\theta}}^{EB(2)}) - R_{\underline{Q}^*}(\underline{\theta}, \underline{X})$$

$$\le \; n(n+p+1)^{-2} E_{\underline{\theta},\underline{V}} \mathrm{tr}\{[(n+\mathrm{tr}(\underline{Q})/(\min \eta_i)+1)\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}$$

$$-4(n+p+1)\underline{D}^{\underline{S}}[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}] - 2(n+p+1)(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}]\underline{V}^{1/2}\underline{Q}\,\underline{V}^{1/2}\}$$

$$= \; n(n+p+1)^{-1} r_0^{-1} E_{\underline{\theta},\underline{V}} \mathrm{tr}\{[\underline{S}^{-1}\underline{g}(\underline{S})\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1} - 4r_0\underline{D}^{\underline{S}}[\underline{S}\,\underline{g}(\underline{S})\underline{S}^{-1}]$$

$$-2r_0(K-p-1)\underline{g}(\underline{S})\underline{S}^{-1}]\underline{V}^{1/2}\underline{Q}\,\underline{V}^{1/2}\}, \qquad (3.4.29)$$

where $r_0 = (n+p+1)/[n+\mathrm{tr}(\underline{Q})/(\min \eta_i)+1]$

As before, we discuss the particular choice of $\underline{g}(\underline{S}) = \underline{P}\,\underline{\Psi}(\underline{\lambda})\,\underline{P}^T$. Then Zheng's (1986) Theorem 2 is extended in the next theorem to the situation that $\underline{V}$ is complete unknown and the loss function is $L_{\underline{Q}^*}$.

Theorem 3.4.2 Estimators of the form $\hat{\underline{\theta}}^{EB(2)}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - \hat{\underline{V}}\,\underline{S}^{-1}\underline{P}\,\underline{\Psi}(\underline{\lambda})\,\underline{P}^T)\underline{X}$ dominate $\underline{X}$ under the $L_{\underline{Q}^*}$ loss if

(i) $0 < \Psi_i(\underline{\lambda}) < 2r_0(K-p-1)$ for each $i = 1, \ldots, p$;

(ii) $\partial\Psi_i(\underline{\lambda})/\partial\lambda_i \ge 0$ for each $i = 1, \ldots, p$;

(iii) $\Psi_i(\underline{\lambda})$'s are similarly ordered with the λ_i's, that is,

$(\Psi_i(\underline{\lambda}) - \Psi_j(\underline{\lambda}))(\lambda_i - \lambda_j) \ge 0$ for every $1 \le i, j \le p$,

where $r_0 = (n+p+1)/[n+\mathrm{tr}(\underline{Q})/(\min \eta_i)+1]$.

Proof of Theorem 3.4.2 See the proof of Theorem 3.3.2

Remark 3.4.2 All the results in Remark 3.3.3 still hold for $\hat{\underline{\theta}}^{EB(2)}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - \hat{\underline{V}}\,\underline{S}^{-1}\underline{P}\,\underline{\Psi}(\underline{L})\underline{P}^T)\underline{X}$ by the obvious replacement of new $r_0 = (n+p+1)/[n+tr(\underline{Q})/(\min \eta_i)+1]$ and $\hat{\underline{V}}_c$ with $\hat{\underline{V}}$

The improvement of $\hat{\underline{\theta}}^{EB(3)}(\hat{\underline{V}}, \underline{X})$ defined in (3.1.13) over $\hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}, \underline{X})$ of the form (3.1.11) is considered in the rest of this section. First, Theorem 3.1.2 is extended based on the assumptions $\hat{\underline{V}} \sim W(p, n, (n+p+1)^{-1}\underline{V})$ and $r(\hat{\underline{V}}, \underline{S}) = a$ is constant.

It follows from (3.1.14) and (3.4.27), the risk difference of $\hat{\underline{\theta}}^{EB(3)}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - a\hat{\underline{V}}\,\underline{S}^{-1} - \hat{\underline{V}}\underline{S}^{-1}\underline{f}(\underline{S}))\underline{X}$ and $\hat{\underline{\theta}}^{EB(1)}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - a\hat{\underline{V}}\,\underline{S}^{-1})\underline{X}$ under the L_{Q^*} loss is

$$n(n+p+1)^{-2}E_{\underline{\theta}\,\underline{V}}tr\{\underline{P}[\underline{\Psi}^2(\underline{L})\underline{\Delta}^{-1}+2a\underline{\Psi}(\underline{L})\underline{\Delta}^{-1}]\underline{P}^T[(n+1)\underline{V}\,\underline{Q}^*\underline{V}+tr(\underline{V}\,\underline{Q}^*)\underline{V}]$$

$$-(n+p+1)\underline{P}[4\underline{\Psi}^{(1)}(\underline{L})+2(K-p-1)\underline{\Psi}(\underline{L})\underline{\Delta}^{-1}]\underline{P}^T\underline{V}\,\underline{Q}^*\underline{V}\}$$

$$\le n(n+p+1)^{-1}r_0^{-1}E_{\underline{\theta}\,\underline{V}}tr\{[\underline{\Psi}^2(\underline{L})\underline{\Delta}^{-1}-2(r_0(K-p-1)-a)\underline{\Psi}(\underline{L})\underline{\Delta}^{-1}-4r_0\underline{\Psi}(\underline{L})\underline{\Delta}^{-1}]$$

$$\underline{P}^T\underline{V}^{1/2}\underline{Q}\,\underline{V}^{1/2}\underline{P}\} \qquad (3.4.30)$$

To show $(\underline{I}_p - a\hat{\underline{V}}\,\underline{S}^{-1} - \hat{\underline{V}}\underline{S}^{-1}\underline{f}(\underline{S}))\underline{X}$ dominates $(\underline{I}_p - a\hat{\underline{V}}\,\underline{S}^{-1})\underline{X}$ under the L_{Q^*} loss is suffices to show (3.4.30) < 0 or $\Delta_{3i}(\underline{L}) < 0$ for each $i = 1, \ldots, p$, where

$$\Delta_{3i}(\underline{L}) = \Psi_i^2(\underline{L})\lambda_i^{-1} - 2[r_0(K-p-1)-a]\Psi_i(\underline{L})\lambda_i^{-1} - 2r_0\sum_{j(\neq i)}[\Psi_i(\underline{L})-\Psi_j(\underline{L})]/(\lambda_i-\lambda_j)$$

$$-4r_0[\partial\Psi_i(\underline{L})/\partial\lambda_i] \qquad (3.4.31)$$

Then the following theorem provides solutions to $\Delta_{3i}(\underline{L}) < 0$ for each $i = 1, \ldots, p$

Theorem 3.4.3 Let $0 < a < r_0(K-2)$ where $r_0 = (n+p+1)/[n+\mathrm{tr}(\underline{Q})/(\min \eta_i)+1]$. Suppose $\Psi_i(\underline{\lambda}) = \lambda_i \tau(U)/(\sum_{j=1}^{p}\lambda_j)$, $i = 1, \ldots, p$, where U is a function of the λ_i's. $(\underline{I}_p - a\hat{\underline{V}}\,\underline{S}^{-1} - [\tau(U)/\mathrm{tr}(\underline{S})]\hat{\underline{V}})\underline{X}$ dominates $(\underline{I}_p - a\hat{\underline{V}}\,\underline{S}^{-1})\underline{X}$ under the L_{Q^*} loss if

(i) $U(\underline{\lambda})$ is non-decreasing in λ_i for each $i = 1, \ldots, p$;

(ii) $\tau(U)$ is non-decreasing in U,

(iii) $0 < \tau(U) < 2[r_0(K-2)-a]$.

Proof of Theorem 3.4.3 See (3.3.27) and (3.3.28) in the proof of Theorem 3.3.4.

Remark 3.4.3 In view of Remark 3.4.1, the matrix extension of the James-Stein estimator $(\underline{I}_p - r_0(K-p-1)\hat{\underline{V}}\,\underline{S}^{-1})\underline{X}$ is the best estimator of the form $(\underline{I}_p - a\hat{\underline{V}}\,\underline{S}^{-1})\underline{X}$, although $(\underline{I}_p - r_0(K-p-1)\hat{\underline{V}}\,\underline{S}^{-1} - [\tau(U)/\mathrm{tr}(\underline{S})]\hat{\underline{V}})\underline{X}$ dominates $(\underline{I}_p - r_0(K-p-1)\hat{\underline{V}}\,\underline{S}^{-1})\underline{X}$ under the L_{Q^*} loss if $0 < \tau(U) < 2r_0(p-1)$. If $\tau(U) = b$ is another constant, then the optimal choice of b is $r_0(p-1)$.

3.5 Simulation Results

To illustrate the the estimation results derived in the previous sections, we compare the simulated risks of five different estimators of $\underline{\theta}$ under the three different assumptions of $\underline{V}$ given in (3.1.8)-(3.1.10). These estimators are $\underline{X}$ and $\hat{\underline{\theta}}_{(j)}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - \hat{\underline{V}}\,\underline{S}^{-1}\underline{r}_{(j)}(\hat{\underline{V}}, \underline{S}))\underline{X}$ ($j = 1, 2, 3, 4$) where

$$\underline{r}_{(1)}(\hat{\underline{V}}, \underline{S}) = r_0\{(K-p-1)\underline{I}_p + [(p-1)/\mathrm{tr}(\underline{S})]\underline{S}\}, \tag{3.5.1}$$

$$\underline{r}_{(2)}(\hat{\underline{V}}, \underline{S}) = r_0\{(K-p-1)\underline{I}_p + [(p-1)/\mathrm{tr}(\underline{S}^2)]\underline{S}^2\}, \tag{3.5.2}$$

$$\underline{r}_{(3)}(\hat{\underline{V}}, \underline{S}) = r_0\{(K-p-1)\underline{I}_p\}, \tag{3.5.3}$$

$$\Sigma_{(4)}(\hat{\underline{V}}, \underline{S}) = (1-\alpha_0)\{(K-p-1)\underline{I}_p+[(p-1)/\mathrm{tr}(\underline{S})]\underline{S}\}+\alpha_0\{[(Kp-2)/\mathrm{tr}(\underline{S})]\underline{S}\}, \qquad (3.5.4)$$

where $\alpha_0 = (K-p-1)/[K^2(p-1)]$ and $r_0 = 1$ under the assumptions of (3.1.8) and (3.1.9), and $\alpha_0 = (p-1)(K-p-1)/[(K-2)-(Kp-2)r_0^{-1}]^2$ and $r_0 = (n+p+1)/[n+\mathrm{tr}(\underline{Q})/(\min \eta_i)+1]$ under the assumption of (3.1.10).

Throught, we have considered the loss (3.1.1). As in Section 2.4 the (K, p) combinations chosen are (4, 2), (8, 2), (12, 2), (5, 3), (9, 3) and (13, 3). For each combination, 1000 samples are generated from $N_p(\underline{\theta}_i, \underline{V})$. Also 1000 samples of $\hat{\underline{V}}$ are generated independently according to the assumption of $\underline{V}$. The constant n, degrees of freedom of the chi-square distribution and the Wishart distribution has been chosen as 3 times K. Also, it is assumed that $\hat{v} \sim v(n+2)^{-1}\chi_n^2$ with v = 1 in the first case.

For p = 2, we have taken $\underline{\theta}_i = \underline{0}$ or $\underline{\theta}_i = (10, 35)^T$ for each i = 1, ..., K. For each p = 3, our choice is $\underline{\theta}_i = \underline{0}$ or $\underline{\theta}_i = (6, 10, 50)^T$ for each i = 1, ..., K. Corresponding to a given situation, for each (K, 2) pair, there are four $(\underline{Q}, \underline{V})$ combinations with each $\underline{Q}$ or $\underline{V}$ equal to $\underline{I}_2$ or $\underline{M}_2 = \mathrm{diag}(0.01, 1.00)$. For each (K, 3) pair, there are once again four $(\underline{Q}, \underline{V})$ combinations with each $\underline{Q}$ or $\underline{V}$ equal to $\underline{I}_3$ or $\underline{M}_3 = \mathrm{diag}(0.01, 1.00, 100.00)$. Thus, in this situation, for each $(K, p, \underline{Q}, \underline{V}, \underline{\theta}_i)$, the simulated risk is the average of 1000 expected losses. There are four tables for each of the three situations of covariance matrix. In each of them, the first two tables report the simulated risks for $\underline{\theta}_i = \underline{0}$ (p×1) while the last two report the same when $\underline{\theta}_i = (10, 35)^T$ and $\underline{\theta}_i = (6, 10, 50)^T$, respectively.

An examination of Tables 3.1-3.12 reveals that like the usual James-Stein estimator, shrinkage estimators of matrix normal means perform the best when the true $\underline{\theta}_i$ equals the point towards which we want to shrink (in this case $\underline{0}$ (p×K)). As the true $\underline{\theta}_i$ gets further and further away from the shrinkage point, the risk improvement also becomes less and less.

Notice that the risk improvement of each estimator keeps nearly the same as the assumption of covariance matrix goes from (3.1.8) to (3.1.9), and decreases as the assumption goes from (3.1.9) to (3.1.10), while all the other conditions are held fixed. Another point to note is that for fixed p, $\underline{Q}$, and $\underline{V}$, the risk improvement of each estimator increases over K. The improvement from K = p+2 to K = p+6 is much more pronounced than the one from K = p+6 to K = p+10. Also, it may be noted that $\hat{\underline{\theta}}_{(1)}$, $\hat{\underline{\theta}}_{(2)}$ and $\hat{\underline{\theta}}_{(4)}$ all outperform $\hat{\underline{\theta}}_{(3)}$ in terms of risk improvement when the true $\underline{\theta}_i$ values are zeros, although for $\underline{\theta}_i$'s far away from zeros, all the four estimators are nearly at par. Finally, we observe that $\hat{\underline{\theta}}_{(4)}$ has a consistent edge over the remaining estimators when $\underline{\theta}_i$'s are zeros, while $\hat{\underline{\theta}}_{(4)}$ performs as well as the others when $\underline{\theta}_i$'s are far away from zeros. Thus, again, we recommend the use of $\hat{\underline{\theta}}_{(4)}$ in the present context.

TABLE 3.1

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_i = \underline{0}$, $\underline{V}$ known up to a positive constant)

	K	$\underline{X}$	RISKS OF $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	PERCENTAGE RISK IMPROVEMENTS BY $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_2$	4	8.0310	5.3456	5.2060	6.3296	5.0551	33.44	35.18	21.19	37.06
	8	16.0644	6.2998	6.0882	6.8076	5.7947	60.78	62.10	57.62	63.93
$\underline{Q} = \underline{I}_2$	12	24.0383	6.8189	6.6296	7.0926	6.4043	71.63	72.42	70.46	73.36
$\underline{V} = \underline{I}_2$	4	4.0211	3.3463	3.2720	3.8418	3.1173	16.78	18.63	4.46	22.48
	8	8.0844	3.2547	3.2437	3.6034	3.0778	58.50	59.88	55.43	61.93
$\underline{Q} = \underline{M}_2$	12	12.1285	3.4552	3.3588	3.5963	3.2421	71.51	72.31	70.35	73.27
$\underline{V} = \underline{M}_2$	4	7.9505	5.6104	5.6494	6.0200	5.6113	29.43	28.95	24.28	29.42
	8	16.0520	6.6623	6.6760	6.8055	6.6573	58.50	58.41	57.60	58.53
$\underline{Q} = \underline{I}_2$	12	24.3445	6.7979	6.8032	6.9249	6.7679	72.08	72.05	71.55	72.20
$\underline{V} = \underline{M}_2$	4	4.2257	2.6423	2.6547	3.2006	2.5782	37.47	37.18	24.26	38.99
	8	7.9605	2.9988	3.0026	3.2194	2.8873	62.33	62.28	59.56	63.73
$\underline{Q} = \underline{M}_2$	12	12.2054	3.6346	3.6385	3.6975	3.6060	70.22	70.19	69.71	70.46

TABLE 3.2

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$ $\underline{\theta}_1 = \underline{0}$, $\underline{V}$ known up to a positive constant)

	K	RISKS OF X	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	PERCENTAGE RISK IMPROVEMENTS BY $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = I_3$	5	14.8437	9.5142	9.2679	11.9814	9.2665	35.90	37.56	19.28	37.57
	9	27.0455	12.2218	11.6955	13.5550	11.7200	54.81	56.76	49.88	56.67
$\underline{Q} = I_3$	13	38.7371	12.4300	11.9124	13.3371	11.9746	67.91	69.25	65.57	69.09
$\underline{V} = I_3$	5	502.9021	317.4052	308.8763	401.3186	309.2105	36.89	38.58	20.20	38.52
	9	903.4710	394.6428	377.7853	439.0219	378.5142	56.32	58.19	51.41	58.10
$\underline{Q} = \underline{M}_3$	13	1303.8896	426.8630	409.8614	456.8245	411.3678	67.26	68.57	64.97	68.45
$\underline{V} = \underline{M}_3$	5	15.0482	10.5936	10.6434	12.0054	10.5588	29.60	29.27	20.22	29.83
	9	27.2885	12.3085	12.3286	13.0716	12.2016	54.89	54.82	52.10	55.29
$\underline{Q} = I_3$	13	38.8555	12.8873	12.9012	13.3016	12.8151	66.83	66.80	65.77	67.02
$\underline{V} = \underline{M}_3$	5	496.9691	260.5431	283.0763	405.6446	277.8731	43.55	43.04	18.38	44.09
	9	919.1461	352.5578	352.7977	422.6580	342.9980	61.64	61.62	54.02	62.68
$\underline{Q} = \underline{M}_3$	13	1307.4704	402.2538	402.7747	437.9532	397.5110	69.23	69.19	68.50	69.60

TABLE 3.3

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($P = 2$, $\underline{\theta}_1 = (10, 35)^T$, $\underline{V}$ known up to a positive constant)

		RISKS OF					PERCENTAGE RISK IMPROVEMENTS BY			
	K	$\underline{X}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = I_2$	4	7.7931	6.8633	6.8639	6.8646	6.8752	11.93	11.92	11.91	11.78
	8	15.8455	11.2869	11.2872	11.2880	11.3025	28.77	28.77	28.76	28.67
$\underline{Q} = I_2$	12	23.7071	15.2780	15.2782	15.2787	15.2896	35.56	35.56	35.55	35.51
$\underline{V} = I_2$	4	4.0190	3.9407	3.9408	3.9427	3.9412	1.95	1.95	1.90	1.94
	8	8.1803	7.7725	7.7726	7.7724	7.7760	4.98	4.98	4.99	4.94
$\underline{Q} = M_2$	12	12.3525	11.6074	11.6074	11.6078	11.6098	6.03	6.03	6.03	6.01
$\underline{V} = M_2$	4	7.9528	7.1272	7.1273	7.1282	7.1267	10.38	10.38	10.37	10.39
	8	16.0645	11.4209	11.4209	11.4215	11.4505	28.91	28.91	28.90	28.72
$\underline{Q} = I_2$	12	24.1010	15.5251	15.5251	15.5250	15.5567	35.58	35.58	35.58	35.45
$\underline{V} = M_2$	4	4.2374	3.4168	3.4169	3.4162	3.4262	19.37	19.36	19.38	19.14
	8	8.0039	3.8895	3.8895	3.8895	3.9181	51.41	51.40	51.40	51.05
$\underline{Q} = M_2$	12	11.9477	4.4816	4.4817	4.4810	4.4964	62.49	62.49	62.49	62.37

TABLE 3.4

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_1 = (6, 10, 50)^T$, $\underline{V}$ known up to a positive constant)

	K	RISKS OF					PERCENTAGE RISK IMPROVEMENTS BY			
		$\underline{X}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_3$	5	14.8760	13.2081	13.2097	13.2114	13.2044	11.21	11.20	11.19	11.23
	9	26.8544	17.2650	17.2659	17.2670	17.2859	35.71	35.71	35.70	35.63
$\underline{Q} = \underline{I}_3$	13	38.9974	21.7425	21.7432	21.7421	21.7614	44.25	44.24	44.25	44.20
$\underline{V} = \underline{I}_3$	5	517.7894	510.9352	510.9394	511.2652	510.9420	1.32	1.32	1.26	1.32
	9	925.8428	897.9482	897.9512	898.0929	997.8700	3.01	3.01	3.00	3.02
$\underline{Q} = \underline{M}_3$	13	1290.4269	1238.4786	1238.4807	1238.5167	1238.4703	4.03	4.03	4.02	4.03
$\underline{V} = \underline{M}_3$	5	15.1530	13.7879	13.7929	13.8141	13.7713	9.01	8.98	8.84	9.12
	9	27.1086	17.4830	17.4857	17.5071	17.5009	35.51	35.50	35.42	35.44
$\underline{Q} = \underline{I}_3$	13	39.0837	21.5868	21.5889	21.5985	21.6153	44.77	44.76	44.74	44.70
$\underline{V} = \underline{M}_3$	5	520.6023	409.6609	410.0846	414.1553	410.1565	21.31	21.23	20.47	21.22
	9	909.4376	425.5395	425.7715	427.1798	426.5070	53.21	53.18	53.03	53.10
$\underline{Q} = \underline{M}_3$	13	1329.4855	458.1635	458.3299	457.4082	457.2269	65.69	65.68	65.60	65.61

TABLE 3.5

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_i = \underline{0}$, $\underline{V}$ unknown diagonal matrix)

	K	$\underline{X}$	RISKS OF $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	PERCENTAGE RISK IMPROVEMENTS BY $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_2$	4	8.0310	5.3118	5.1761	6.3054	5.0224	33.88	35.55	21.49	37.46
	8	18.0644	6.3210	6.1038	6.8140	5.8162	60.85	62.01	57.58	63.79
$\underline{Q} = \underline{I}_2$	12	24.0383	6.7203	8.5354	7.0020	6.3135	72.04	72.81	70.87	73.74
$\underline{V} = \underline{I}_2$	4	4.0211	2.9439	2.8692	3.4319	2.7676	26.79	28.65	14.65	31.17
	8	8.0844	3.5017	3.3841	3.7298	3.2126	56.89	58.14	53.86	60.26
$\underline{Q} = \underline{M}_2$	12	12.1285	3.4552	3.3573	3.5969	3.2419	71.51	72.32	70.34	73.27
$\underline{V} = \underline{M}_2$	4	7.9505	5.8483	5.8932	6.1618	5.8680	26.44	25.88	22.50	28.20
	8	16.0520	6.8220	6.8373	6.9008	6.8603	57.50	57.41	57.01	57.28
$\underline{Q} = \underline{I}_2$	12	24.3445	6.9044	6.9100	7.0178	6.8821	71.64	71.62	71.17	71.73
$\underline{V} = \underline{M}_2$	4	4.2257	2.7173	2.7251	3.2311	2.6673	35.70	35.51	23.54	36.88
	8	7.9605	3.2104	3.2156	3.3632	3.1382	59.67	59.61	57.75	60.58
$\underline{Q} = \underline{M}_2$	12	12.2054	3.5845	3.5873	3.6597	3.5453	70.63	70.61	70.02	70.95

TABLE 3.6

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_i = \underline{0}$, $\underline{V}$ unknown diagonal matrix)

		RISKS OF					PERCENTAGE RISK IMPROVEMENTS BY			
	K	$\underline{X}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = I_3$	5	14.8437	9.5395	9.2960	11.9618	9.2988	35.73	37.37	19.41	37.36
	9	27.0455	12.1328	11.5982	13.4593	11.6317	55.14	57.12	50.23	56.99
$\underline{Q} = I_3$	13	38.7371	12.5206	12.0014	13.4368	12.0597	67.68	69.02	65.31	68.87
$\underline{V} = I_3$	5	502.9061	323.6459	315.1213	406.4415	315.3923	35.64	37.34	19.18	37.29
	9	903.4710	400.0296	382.8745	443.9327	383.4673	55.72	57.62	50.86	57.56
$\underline{Q} = \underline{M}_3$	13	1303.8896	438.2798	421.2074	468.7850	422.3580	66.39	67.70	64.55	67.61
$\underline{V} = \underline{M}_3$	5	15.0482	10.9326	10.9890	12.3136	10.8889	27.35	26.98	18.17	27.64
	9	27.2885	12.5901	12.6079	13.3204	12.4762	53.86	53.80	51.19	54.28
$\underline{Q} = I_3$	13	38.8555	12.6521	12.6656	13.0997	12.5763	67.44	67.40	66.29	67.63
$\underline{V} = \underline{M}_3$	5	496.9691	326.7783	330.4236	429.3686	326.5323	34.25	33.51	13.60	34.30
	9	919.1461	355.0196	355.2107	426.3373	344.6301	61.38	61.35	53.62	62.51
$\underline{Q} = \underline{M}_3$	13	1307.4704	410.3199	410.8422	446.0470	405.4595	68.62	68.58	65.88	68.99

TABLE 3.7

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($P = 2$, $\underline{\theta}_1 = (10, 35)^T$, $\underline{V}$ unknown diagonal matrix)

	K	$\underline{X}$	RISKS OF $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	PERCENTAGE RISK IMPROVEMENTS BY $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_2$	4	7.7931	6.8734	6.8740	6.8747	6.8848	11.80	11.79	11.79	11.65
	8	15.8455	11.2896	11.2900	11.2906	11.3052	28.75	28.75	28.75	28.65
$\underline{Q} = \underline{I}_2$	12	23.7071	15.2649	15.2652	15.2658	15.2797	35.61	35.61	35.61	35.54
$\underline{V} = \underline{I}_2$	4	4.0190	3.9449	3.9449	3.9472	3.9453	1.84	1.84	1.79	1.83
	8	8.1803	7.8068	7.8068	7.8065	7.8077	4.57	4.57	4.57	4.55
$\underline{Q} = \underline{M}_2$	12	12.3525	11.6067	11.6067	11.6073	11.6112	6.04	6.04	6.03	6.00
$\underline{V} = \underline{M}_2$	4	7.9528	7.0439	7.0440	7.0446	7.0528	11.43	11.43	11.42	11.32
	8	16.0645	11.4331	11.4331	11.4336	11.4605	28.83	28.83	28.83	28.66
$\underline{Q} = \underline{I}_2$	12	24.1010	15.5549	15.5549	15.5545	15.6034	35.46	35.46	35.46	35.26
$\underline{V} = \underline{M}_2$	4	4.2374	3.4906	3.4907	3.4900	3.4915	17.62	17.62	17.64	17.60
	8	8.0039	4.0219	4.0220	4.0221	4.0300	49.75	49.75	49.75	49.64
$\underline{Q} = \underline{M}_2$	12	11.9477	4.4242	4.4243	4.4237	4.4641	62.97	62.97	62.97	62.64

TABLE 3.8

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_i = (6, 10, 50)^T$, $\underline{V}$ unknown diagonal matrix)

	K	RISKS OF $\underline{X}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	PERCENTAGE RISK IMPROVEMENTS BY $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_3$	5	14.8760	13.2247	13.2263	13.2280	13.2209	11.10	11.09	11.08	11.13
	9	26.8544	17.2476	17.2485	17.2497	17.2675	35.77	35.77	35.77	35.70
$\underline{Q} = \underline{I}_3$	13	38.9974	21.6770	21.6777	21.6767	21.7029	44.41	44.41	44.41	44.35
$\underline{V} = \underline{I}_3$	5	517.7894	510.5119	510.5163	510.9077	510.5077	1.41	1.40	1.33	1.41
	9	925.8428	896.5789	896.5820	896.7428	896.5626	3.16	3.16	3.14	3.16
$\underline{Q} = \underline{M}_3$	13	1290.4269	1237.1165	1237.1187	1237.1616	1237.1502	4.13	4.13	4.13	4.13
$\underline{V} = \underline{M}_3$	5	15.1530	13.5407	13.5458	13.5653	13.5347	10.64	10.61	10.48	10.68
	9	27.1086	17.3824	17.3852	17.4062	17.4084	35.88	35.87	35.80	35.78
$\underline{Q} = \underline{I}_3$	13	39.0837	21.5522	21.5543	21.5626	21.5844	44.86	44.85	44.83	44.77
$\underline{V} = \underline{M}_3$	5	520.6023	422.5273	422.9396	426.6082	422.5601	18.84	18.76	18.05	18.83
	9	909.4376	433.4318	433.6681	435.0180	434.4007	52.34	52.31	52.17	52.23
$\underline{Q} = \underline{M}_3$	13	1329.4855	450.2434	450.4104	451.9519	451.1204	66.13	66.12	66.01	66.07

TABLE 3.9

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 2$, $\underline{\theta}_i = \underline{0}$, $\underline{V}$ completely unknown)

	K	$\underline{X}$	RISKS OF				PERCENTAGE RISK IMPROVEMENTS BY			
			$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_2$	4	8.0310	5.4576	5.3339	6.3985	5.1867	32.04	33.58	20.33	35.42
	8	16.0644	8.6957	6.4891	7.1687	6.2097	58.32	59.61	55.38	61.34
$\underline{Q} = \underline{I}_2$	12	24.0383	7.2207	7.0392	7.4921	6.8216	69.96	70.72	68.83	71.62
$\underline{V} = \underline{I}_2$	4	4.0211	3.7267	3.7249	3.8270	3.0212	7.32	7.37	4.83	24.87
	8	8.0844	6.2105	6.2049	6.3728	3.6093	23.18	23.25	21.17	55.36
$\underline{Q} = \underline{M}_2$	12	12.1285	7.8552	7.8486	8.0537	3.6995	35.23	35.29	33.60	69.50
$\underline{V} = \underline{M}_2$	4	7.9505	6.0021	6.0423	6.3057	6.0129	24.51	24.00	20.69	24.37
	8	16.0520	7.1546	7.1696	7.2239	7.2005	55.43	55.34	55.00	55.14
$\underline{Q} = \underline{I}_2$	12	24.3445	7.3446	7.3501	7.4566	7.3216	69.83	69.81	69.37	69.93
$\underline{V} = \underline{M}_2$	4	4.2257	3.8354	3.8334	4.0248	2.7366	9.24	9.28	4.75	35.24
	8	7.9605	5.9357	5.9326	6.2416	3.3422	25.44	25.47	21.59	58.01
$\underline{Q} = \underline{M}_2$	12	12.2054	7.7451	7.7413	8.1256	3.7352	36.54	36.57	33.43	69.40

TABLE 3.10

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_i = \underline{0}$, $\underline{V}$ completely unknown)

	K	RISKS OF $\underline{X}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	PERCENTAGE RISK IMPROVEMENTS BY $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_3$	5	14.8437	10.1074	9.8873	12.2720	9.8922	31.91	33.39	17.33	33.36
	9	27.0455	13.2565	12.7477	14.4949	12.7789	50.98	52.87	46.41	52.75
$\underline{Q} = \underline{I}_3$	13	38.7371	13.6275	13.1357	14.5070	13.1911	64.82	66.09	62.55	65.95
$\underline{V} = \underline{I}_3$	5	502.9061	502.4109	502.4105	502.6080	345.4867	10	10	06	31.30
	9	903.4710	900.4312	900.4342	900.7827	432.5861	34	34	30	52.12
$\underline{Q} = \underline{M}_3$	13	1303.8896	1296.4862	1296.4867	1296.9937	473.1737	57	57	53	63.71
$\underline{V} = \underline{M}_3$	5	15.0482	11.6504	11.7004	12.8669	11.6032	22.58	22.25	14.50	22.89
	9	27.2885	13.6260	13.6427	14.2838	13.5250	50.07	50.01	47.66	50.44
$\underline{Q} = \underline{I}_3$	13	38.8555	14.0560	14.0689	14.4651	13.9831	63.83	63.79	62.77	64.01
$\underline{V} = \underline{M}_3$	5	496.9691	496.0902	496.0833	496.6710	322.0777	18	18	06	35.19
	9	919.1461	915.4069	915.3957	916.4575	379.7203	41	41	29	58.69
$\underline{Q} = \underline{M}_3$	13	1307.4704	1299.0775	1299.0612	1300.5742	441.3433	64	64	53	66.24

TABLE 3.11

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ (P = 2, $\underline{\theta}_1 = (10, 35)^T$, $\underline{V}$ completely unknown)

	K	X	RISKS OF $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	PERCENTAGE RISK IMPROVEMENTS BY $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_2$	4	7.7931	6.9171	6.9176	6.9186	6.9284	11.24	11.23	11.22	11.10
	8	15.8455	11.4995	11.4998	11.5003	11.5101	27.43	27.43	27.42	27.36
$\underline{Q} = \underline{I}_2$	12	23.7071	15.3856	15.3859	15.3865	15.4097	35.10	35.10	35.10	35.00
$\underline{V} = \underline{I}_2$	4	4.0190	4.0038	4.0038	4.0041	3.9894	.38	.38	.37	.74
	8	8.1803	8.0397	8.0397	8.0398	7.9291	1.72	1.72	1.72	3.07
$\underline{Q} = \underline{M}_2$	12	12.3525	12.0122	12.0122	12.0126	11.7510	2.75	2.75	2.75	4.87
$\underline{V} = \underline{M}_2$	4	7.9528	7.1111	7.1111	7.1121	7.1183	10.58	10.58	10.57	10.49
	8	16.0645	11.6284	11.6284	11.6290	11.6513	27.61	27.61	27.61	27.47
$\underline{Q} = \underline{I}_2$	12	24.1010	15.7316	15.7317	15.7314	15.7845	34.73	34.73	34.73	34.51
$\underline{V} = \underline{M}_2$	4	4.2374	4.0604	4.0604	4.0603	3.4978	4.18	4.18	4.18	17.46
	8	8.0039	6.4944	6.4944	6.4946	4.0353	18.86	18.86	18.86	49.58
$\underline{Q} = \underline{M}_2$	12	11.9477	8.4102	8.4103	8.4103	4.4615	29.61	29.61	29.61	62.66

TABLE 3.12

THE RISKS AND PERCENTAGE RISK IMPROVEMENTS OVER $\underline{X}$ ($p = 3$, $\underline{\theta}_1 = (6, 10, 50)^T$, $\underline{V}$ completely unknown)

	K	RISKS OF $\underline{X}$	$\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$	PERCENTAGE RISK IMPROVEMENTS BY $\hat{\underline{\theta}}_{(1)}$	$\hat{\underline{\theta}}_{(2)}$	$\hat{\underline{\theta}}_{(3)}$	$\hat{\underline{\theta}}_{(4)}$
$\underline{V} = \underline{I}_3$	5	14.8760	13.4112	13.4126	13.4140	13.4075	9.85	9.84	9.83	9.87
	9	26.8546	17.8395	17.8404	17.8416	17.8592	33.57	33.57	33.56	33.50
$\underline{Q} = \underline{I}_3$	13	38.9974	22.4613	22.4620	22.4611	22.4859	42.40	42.40	42.40	42.34
$\underline{V} = \underline{I}_3$	5	517.7894	517.7719	517.7719	517.7727	519.6580	.00	.00	.00	.36
	9	925.8428	925.6801	925.6801	925.6809	926.9266	.02	.02	.02	-.12
$\underline{Q} = \underline{M}_3$	13	1290.4269	1289.9662	1289.9662	1289.9668	1267.9314	.04	.04	.04	1.74
$\underline{V} = \underline{M}_3$	5	15.1530	13.6918	13.6963	13.7182	13.6865	9.64	9.61	9.47	9.68
	9	27.1086	18.0309	18.0335	18.0528	18.0554	33.49	33.48	33.41	33.40
$\underline{Q} = \underline{I}_3$	13	39.0837	22.4526	22.4546	22.4656	22.4782	42.55	42.55	42.52	42.49
$\underline{V} = \underline{M}_3$	5	520.6023	520.2878	520.2887	520.3034	428.1523	.06	.06	.06	17.76
	9	909.4376	906.7661	906.7678	906.7961	451.4563	.29	.29	.29	50.36
$\underline{Q} = \underline{M}_3$	13	1329.4855	1322.4701	1322.4726	1322.5207	471.1727	.53	.53	.52	64.56

CHAPTER 4
EMPIRICAL BAYES MINIMAX ESTIMATORS OF MATRIX NORMAL MEANS FOR ARBITRARY QUADRATIC LOSS AND KNOWN COVARIANCE MATRIX IN THE TWO-SAMPLE PROBLEM

4.1 Introduction

We consider the problem of estimation of one of the two sets of matrix normal means when it is suspected that the two matrix normal means are equal. Let $\underline{X}_{1i}$ and $\underline{X}_{2i}$ be independent p-dimensional column vectors ($i = 1, \ldots, K$, $K > p+1$), where $\underline{X}_{1i}$'s are independently distributed as $N_p(\underline{\theta}_{1i}, \underline{V})$, while $\underline{X}_{2i}$'s are independently distributed as $N_p(\underline{\theta}_{2i}, \underline{V})$, that is,

$$\underline{X}_{hi} \mid \underline{\theta}_{hi} \sim N_p(\underline{\theta}_{hi}, \underline{V}), \tag{4.1.1}$$

$h = 1, 2$, $i = 1, \ldots, K$. In the above, $\underline{\theta}_{hi}$'s are unknown, but $\underline{V}$ is a known p×p symmetric p.d. matrix. Let $\underline{\theta}_h = (\underline{\theta}_{h1}, \ldots, \underline{\theta}_{hK})$, $X_h = (\underline{X}_{h1}, \ldots, \underline{X}_{hK})$, $h = 1, 2$. Our goal is to estimate $\underline{\theta}_1$. We will consider the same problem with an unknown covariance matrix $\underline{V}$ in Chapter 5. Such problem can appear in a balanced two-way MANOVA model

$$\underline{X}_{hij} = \underline{\theta}_{hi} + \underline{\varepsilon}_{hij}, \tag{4.1.2}$$

$h = 1, 2$, $i = 1, \ldots, K$, $j = 1, \ldots, m$, where $\underline{X}_{hij}$ is a p×1 vector of responses of the jth individual in hth block receiving the ith treatment, and $\underline{\theta}_{hi}$ is the p×1 vector of ith treatment and hth block effect combination. The vectors $\underline{\varepsilon}_{hij}$'s are independently distributed as $N_p(\underline{0}, \underline{V}_0)$, where $\underline{V}_0$ is a known p×p symmetric p.d. matrix. The minimal sufficient statistic for $\underline{\theta}_h = (\underline{\theta}_{h1}, \ldots, \underline{\theta}_{hK})$ is $X_h = (\underline{X}_{h1}, \ldots, \underline{X}_{hK})$, where $\underline{X}_{hi} = m^{-1}\sum_{j=1}^{m} \underline{X}_{hij}$, $h = 1, 2$. The $\underline{X}_{hi}$'s are now independently distributed as $N_p(\underline{\theta}_{hi}, \underline{V})$, where $\underline{V} = m^{-1}\underline{V}_0$ and $K >$

p+1 It is also suspected that there is no difference in block effects ($\underline{\theta}_{1i} = \underline{\theta}_{2i}$, $i = 1, \ldots, K$), in which case, observations in second block can possibly be used effectively together with those in first block for estimating the population means of the first block. Thus, the question that naturally arises is whether one should use the sample mean $\underline{X}_{11}$ from block one or the pooled sample mean $(\underline{X}_{11}+\underline{X}_{21})/2$ from both blocks One natural way to settle this problem is to test the null hypothesis of the equality of the two block effects, and then use a compromise estimator which leans more towards the sample mean from block one if the null hypothesis is rejected, and use the pooled sample mean when such a hypothesis is accepted. A very popular method to achieve this compromise is to use a preliminary test estimator (PTE) which uses the pooled sample mean when the null hypothesis is accepted at a desired level of significance and uses the sample mean from block one if the opposite is true The above procedure suffers from the drawback that there is no general way of incorporating the degree of evidence for or against the null hypothesis in order to arrive at the estimator

In this chapter we propose instead three classes of EB estimators which achieve the intended compromise. Such an EB estimator is quite often a weighted average of the pooled sample mean and the first sample mean The weights are adaptively determined from the data in such a way that the larger the value of the statistic used for testing the equality of the two block effects, the smaller is the weight attached to the pooled sample mean Thus, unlike the PTE, the EB estimator incorporates the degree of evidence for or against the null hypothesis in a very natural way Furthermore, we provide modified versions of the EB estimators which still are weighted averages of the pooled sample mean and the first sample mean but the outcomes are either the pooled sample mean or the corresponding EB estimators depending on the acceptance or rejection of the hypothesis Such estimators continue to incorporate the degree of evidence for or against the null hypothesis in a very natural way and are shown to be better than the PTE in frequentist sense.

In order to motivate the EB estimator, we need a Bayes procedure. The distribution in (4.1.1) is treated as conditional distribution given $\underline{\theta}_{hi}$ in Bayesian framework. Suppose also that $\underline{\theta}_{hi}$'s have independent distributions

$$\underline{\theta}_{hi} \sim N_p(\underline{\mu}_i, \underline{\Delta}), \tag{4.1.3}$$

where $\underline{\mu}_i$'s are p×1 column vectors and $\underline{\Delta}$ is a p×p symmetric p.d. matrix. The suspicion that $\underline{\theta}_{1i}$ and $\underline{\theta}_{2i}$ may be equal is reflected in the choice of a prior common mean $\underline{\mu}_i$. Note that conditional on $\underline{\theta}_{1i}$'s and $\underline{\theta}_{2i}$'s, $\underline{X}_{1i}$'s and $\underline{X}_{2i}$'s are mutually independent. Also, $\underline{\theta}_{1i}$'s and $\underline{\theta}_{2i}$'s have independent normal priors, standard calculations show that $\underline{\theta}_{1i}$'s and $\underline{\theta}_{2i}$'s given $\underline{X}_{1i}$'s and $\underline{X}_{2i}$'s have mutually independent posterior distributions

$$\underline{\theta}_{hi} \mid \underline{X}_{hi} = \underline{x}_{hi} \sim N_p((\underline{I}_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{x}_{hi} + \underline{V}\,\underline{\Sigma}^{-1}\underline{\mu}_i,\ (\underline{I}_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{V}), \tag{4.1.4}$$

where $\underline{\Sigma} = (\underline{V}+\underline{\Delta})$. Now, using the loss function

$$L_Q(\underline{\theta}_1, \underline{a}) = \sum_{i=1}^{K} (\underline{\theta}_{1i} - \underline{a}_i)^T Q(\underline{\theta}_{1i} - \underline{a}_i) \tag{4.1.5}$$

for estimating $\underline{\theta}_1$, where Q is again a known p×p p.d. matrix. The Bayes estimator of $\underline{\theta}_1$ is given by

$$\hat{\underline{\theta}}_1^B(\underline{X}) = (\underline{I}_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{X}_1 + \underline{V}\,\underline{\Sigma}^{-1}\underline{\mu}, \tag{4.1.6}$$

where $\underline{X} = (\underline{X}_1, \underline{X}_2)$ and $\underline{\mu} = (\underline{\mu}_1, \ldots, \underline{\mu}_K)$. In order to find an EB estimator of $\underline{\theta}_1$, we estimate the unknown parameters $\underline{\mu} = (\underline{\mu}_1, \ldots, \underline{\mu}_K)$ and $\underline{\Sigma}$ from the marginal distributions of $\underline{X}_{1i}$'s and $\underline{X}_{2i}$'s. Note that marginally $\underline{X}_{hi}$'s are mutually independent with

$$\underline{X}_{hi} \sim N_p(\underline{\mu}_i, \underline{\Sigma}) \tag{4.1.7}$$

Then $(\underline{Z}, \underline{S})$ is complete sufficient for $(\underline{\mu}, \underline{\Sigma})$ base on the marginal distributions where $\underline{Z} = (\underline{Z}_1, \ldots, \underline{Z}_K)$, $\underline{Z}_i = (\underline{X}_{1i}+\underline{X}_{2i})/2 \sim N_p(\underline{\mu}_i, 2^{-1}\underline{\Sigma})$, $\underline{Y}_i = 2(\underline{X}_{1i} - \underline{Z}_i) \sim N_p(\underline{0}, 2\underline{\Sigma})$ and $\underline{S} =$

$2^{-1}(\sum_{i=1}^{K}\underline{Y}_i\underline{Y}_i^T) \sim W(p, K, \underline{\Sigma})$. Hence the UMVUE of $\underline{\mu}$ is $\underline{Z}$ and the best multiple estimator of $\underline{\Sigma}^{-1}$ is $(K-p-1)\underline{S}^{-1}$. Substituting these estimators for $\underline{\mu}$ and $\underline{\Sigma}^{-1}$ in (4.1.6), one gets the EB estimator $(\underline{I}_p - (K-p-1)\underline{V}\,\underline{S}^{-1})\underline{X}_1 + (K-p-1)\underline{V}\,\underline{S}^{-1}\underline{Z}$ of $\underline{\theta}_1$. A general class of EB estimators of $\underline{\theta}_1$ is given by

$$\hat{\underline{\theta}}_1^{EB}(\underline{X}) = (\underline{I}_p - \underline{V}\,\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S}))\underline{X}_1 + \underline{V}\,\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S})\underline{Z}, \qquad (4.1.8)$$

where $\underline{\tau}(\underline{V}, \underline{S})$ is a $p \times p$ symmetric matrix. As in Chapter 2, we prefer this notation even if $\underline{V}$ is known, since later in the case of unknown $\underline{V}$, we will use $\underline{\tau}(\hat{\underline{V}}, \underline{S})$, where $\hat{\underline{V}}$ estimates $\underline{V}$.

In Section 4.2, we consider three classes of EB estimators which stem from (4.1.8). It is well known that under the loss given in (4.1.5), $\underline{X}_1$ is a minimax estimator of $\underline{\theta}_1$ with constant risk $K\mathrm{tr}(\underline{V}\,\underline{Q})$. We first find the risk difference of $\hat{\underline{\theta}}_1^{EB}$ and $\underline{X}_1$, and then outline the methods of finding EB minimax estimators in Section 2.2 and 2.3 without repeating the materials of Chapter 2. Section 4.3 provides modified versions of these EB minimax estimators as proposed in Section 4.2, and later are shown to dominate the PTE.

The ideas of this chapter are taken from Ghosh and Sinha (1988) who considered the problem of estimating a vector of normal means in the two-sample situation and proposed both EB and HB estimators which dominated the first sample mean.

The problem can be extended to multi-sample case ($h = 1, \ldots, H$, $H > 2$), in which case, several matrix normal means are suspected to be equal, and observations in all the samples can be combined effectively to improve estimation of a matrix of normal means. This is discussed in Remark 4.2.1

4.2. EB Minimax Estimators Under The L_Q Loss

We first compute the risk of the estimator $\hat{\underline{\theta}}_1^{EB}$ as defined in (4.1.8). Throughout this chapter, we denote by $E_{\underline{\theta}}$ the expectation conditional on $\underline{\theta} = (\underline{\theta}_1, \underline{\theta}_2)$, $\tilde{E}$ the expectation over the marginal distribution of $\underline{\theta}$, and E the expectation over the joint distribution of $\underline{X} = (\underline{X}_1, \underline{X}_2)$ and $\underline{\theta}$. Also, let $R_Q(\underline{\theta}_1, \underline{\varrho}(\underline{X})) = E_{\underline{\theta}}[L_Q(\underline{\theta}_1, \underline{\varrho}(\underline{X}))]$, where L_Q is defined in (4.1.5).

Theorem 4.2.1 Under the L_Q loss given in (4.1.5), $\hat{\underline{\theta}}_1^{EB}$ has risk

$$
\begin{aligned}
& R_Q(\underline{\theta}_1, \hat{\underline{\theta}}_1^{EB}) \\
= \; & R_Q(\underline{\theta}_1, \underline{X}_1) + \tfrac{1}{2} E_{\underline{\theta}} \mathrm{tr}\{[\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S})\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1} - 4\underline{D}^{S}[\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}] \\
& - 2(K-p-1)\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}]\underline{Q}_0\}, \qquad (4.2.1)
\end{aligned}
$$

where $R_Q(\underline{\theta}_1, \underline{X}_1) = K\mathrm{tr}(\underline{V}\,\underline{Q})$, $\underline{Q}_0 = \underline{V}\,\underline{Q}\,\underline{V}$ and $\underline{D}^S$ is defined in Lemma 1.3.3.

Proof of Theorem 4.2.1 Let $\hat{\underline{\theta}}_{1i}^{B}(\underline{X}) = (\underline{I}_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{X}_{1i} + \underline{V}\,\underline{\Sigma}^{-1}\underline{\mu}_i$ and $\hat{\underline{\theta}}_{1i}^{EB}(\underline{X}) = (\underline{I}_p - \underline{V}\,\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S}))\underline{X}_{1i} + \underline{V}\,\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S})\underline{Z}_i$, i = 1, ..., K. Standard calculation yields

$$
\begin{aligned}
& EL_Q(\underline{\theta}_1, \hat{\underline{\theta}}_1^{EB}) \\
= \; & EL_Q(\underline{\theta}_1, \hat{\underline{\theta}}_1^{B}) + E\{\sum_{i=1}^{K} (\hat{\underline{\theta}}_{1i}^{B} - \hat{\underline{\theta}}_{1i}^{EB})^T \underline{Q} (\hat{\underline{\theta}}_{1i}^{B} - \hat{\underline{\theta}}_{1i}^{EB})\}, \qquad (4.2.2)
\end{aligned}
$$

where

$$
\begin{aligned}
& \hat{\underline{\theta}}_{1i}^{B} - \hat{\underline{\theta}}_{1i}^{EB} \\
= \; & \underline{V}(\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S}) - \underline{\Sigma}^{-1})(\underline{X}_{1i} - \underline{Z}_i) - \underline{V}\,\underline{\Sigma}^{-1}(\underline{Z}_i - \underline{\mu}_i) \\
= \; & \tfrac{1}{2}\underline{V}(\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S}) - \underline{\Sigma}^{-1})\underline{Y}_i - \underline{V}\,\underline{\Sigma}^{-1}(\underline{Z}_i - \underline{\mu}_i) \qquad (4.2.3)
\end{aligned}
$$

Next use the independence of $\underline{Z}_i$ and $\underline{Y}_i$ (and $\underline{S}$) and the fact $\underline{Z}_i \sim N_p(\underline{\mu}_i, 2^{-1}\Sigma)$, one gets

$$
\begin{aligned}
& E\{\sum_{i=1}^{K}(\hat{\underline{\theta}}_{1i}^{B}-\hat{\underline{\theta}}_{1i}^{EB})^{T}\underline{Q}(\hat{\underline{\theta}}_{1i}^{B}-\hat{\underline{\theta}}_{1i}^{EB})\} \\
= \; & E\{\sum_{i=1}^{K}(\underline{Z}_i-\underline{\mu}_i)^{T}\Sigma^{-1}\underline{V}\,\underline{Q}\,\underline{V}\Sigma^{-1}(\underline{Z}_i-\underline{\mu}_i) \\
& +\frac{1}{4}\sum_{i=1}^{K}\underline{Y}_i^{T}(\underline{S}^{-1}\underline{\tau}(\underline{V},\underline{S})-\Sigma^{-1})^{T}\underline{V}\,\underline{Q}\,\underline{V}(\underline{S}^{-1}\underline{\tau}(\underline{V},\underline{S})-\Sigma^{-1})\underline{Y}_i\} \\
= \; & E\mathrm{tr}\{(\sum_{i=1}^{K}(\underline{Z}_i-\underline{\mu}_i)(\underline{Z}_i-\underline{\mu}_i)^{T})\Sigma^{-1}\underline{Q}_0\Sigma^{-1} \\
& +\frac{1}{2}(\underline{S}^{-1}\underline{\tau}(\underline{V},\underline{S})-\Sigma^{-1})(\frac{1}{2}\sum_{i=1}^{K}\underline{Y}_i\underline{Y}_i^{T})(\underline{\tau}^{T}(\underline{V},\underline{S})\underline{S}^{-1}-\Sigma^{-1})\underline{Q}_0\} \\
= \; & \frac{K}{2}\mathrm{tr}(\Sigma^{-1}\underline{Q}_0)+\frac{1}{2}E\mathrm{tr}\{(\underline{S}^{-1}\underline{\tau}(\underline{V},\underline{S})-\Sigma^{-1})\underline{S}(\underline{\tau}^{T}(\underline{V},\underline{S})\underline{S}^{-1}-\Sigma^{-1})\underline{Q}_0\} \qquad (4.2.4)
\end{aligned}
$$

Applying the same argument to obtain (2.2.4) and the fact $E(\underline{S}) = K\Sigma$, it follows that the second term on the right-hand side of (4.2.4) is

$$\frac{1}{2}E\mathrm{tr}\{[\underline{S}^{-1}\underline{\tau}(\underline{V},\underline{S})\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}-2\Sigma^{-1}\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}]\underline{Q}_0\}+\frac{K}{2}\mathrm{tr}(\Sigma^{-1}\underline{Q}_0) \qquad (4.2.5)$$

From (4.2.4) and (4.2.5), it follows that

$$
\begin{aligned}
& E\{\sum_{i=1}^{K}(\hat{\underline{\theta}}_{1i}^{B}-\hat{\underline{\theta}}_{1i}^{EB})^{T}\underline{Q}(\hat{\underline{\theta}}_{1i}^{B}-\hat{\underline{\theta}}_{1i}^{EB})\} \\
= \; & K\mathrm{tr}(\Sigma^{-1}\underline{Q}_0)+\frac{1}{2}E\mathrm{tr}\{[\underline{S}^{-1}\underline{\tau}(\underline{V},\underline{S})\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}-2\Sigma^{-1}\underline{S}\,\underline{\tau}(\underline{V},\underline{S})\underline{S}^{-1}]\underline{Q}_0\} \qquad (4.2.6)
\end{aligned}
$$

Note from (4.1.4) and (4.1.6) that

$$EL_{\underline{Q}}(\underline{\theta}_1, \hat{\underline{\theta}}_1^{B}) = K\mathrm{tr}\{(\underline{I}_p-\underline{V}\Sigma^{-1})\underline{V}\,\underline{Q}\} \qquad (4.2.7)$$

Thus from (4.2.2), (4.2.6) and (4.2.7),

$$EL_Q(\underline{\theta}_1, \hat{\underline{\theta}}_1^{EB})$$

$$= K\mathrm{tr}(\underline{\Psi}\, Q_0) + \frac{1}{2} E\mathrm{tr}\{[\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S})\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}\ 2\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}]Q_0\} \tag{4.2.8}$$

Next apply the Wishart identity as given in Lemma 1.3.3, and get

$$E\{\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}\} = E\{2\underline{D}^{S}\{\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}\} + (K\ p-1)\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}\} \tag{4.2.9}$$

Note that $R_Q(\underline{\theta}_1, \underline{X}_1) = EL_Q(\underline{\theta}_1, \underline{X}_1) = K\mathrm{tr}(\underline{\Psi}\, Q)$ for all $\underline{\theta}$. Hence we find

$$EL_Q(\underline{\theta}_1, \hat{\underline{\theta}}_1^{EB})$$

$$= \tilde{E}R_Q(\underline{\theta}_1, \hat{\underline{\theta}}_1^{EB})$$

$$= \tilde{E}R_Q(\underline{\theta}_1, \underline{X}_1) + \tilde{E}E_{\underline{\theta}}\mathrm{tr}\{[\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S})\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1} - 4\underline{D}^{S}\{\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}\}$$

$$2(K-p-1)\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}]Q_0\} \tag{4.2.10}$$

The conditional distribution of $\underline{S} = 2^{-1}\sum_{i=1}^{K}\underline{Y}_i\underline{Y}_i^T$ given $\underline{\theta}$ depends on $\underline{\theta}$ only through $2^{-1}\sum_{i=1}^{K}(\underline{\theta}_{1i}\ \underline{\theta}_{2i})(\underline{\theta}_{1i}\ \underline{\theta}_{2i})^T = \underline{\Theta}^*$(say) Also the family of distributions of $\underline{\Theta}$ being Wishart $W(p, K, \underline{\Delta})$ is complete for $\underline{\Delta}$ (or $\underline{\Sigma}$) The proof of the theorem is complete from (4 2 10)

In view of (4.2 1) an unbiased estimator of the risk difference $R_Q(\underline{\theta}_1\ \hat{\underline{\theta}}_1^{EB})$ $R_Q(\underline{\theta}_1, \underline{X}_1)$ is given by

$$\frac{1}{2}\mathrm{tr}\{Q_0^{1/2}[\underline{S}^{-1}\underline{\tau}(\underline{V}, \underline{S})\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}\ 4\underline{D}^{S}[\underline{S}\,\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}]$$

$$2(K\ p-1)\underline{\tau}(\underline{V}, \underline{S})\underline{S}^{-1}]Q_0^{1/2}\}, \tag{4.2.11}$$

where $Q_0 = Q_0^{1/2}Q_0^{1/2}$. If we can find $r(\underline{V}, \underline{S})$ such that (4.2.11) is non-positive every where and is negative with positive probability for all $\underline{\theta}$, then we obtain EB estimator $\hat{\underline{\theta}}_1^{EB}$ which dominates $\underline{X}_1$ under the L_Q loss and is minimax afterward. Notice the analogy between (4.2.11) and the unbiased estimator of the risk difference $R_Q(\underline{\theta}, \hat{\underline{\theta}}^{EB})-R_Q(\underline{\theta}, \underline{X})$ obtained from (2.2.1) in one sample problem. In both cases, $\underline{V}$, $\underline{S}$ and $r(\underline{V}, \underline{S})$ are pxp symmetric p.d. matrices. We can proceed exactly the same way in Section 2.2 to find two classes of EB minimax estimators $\hat{\underline{\theta}}_1^{EB(1)}$ and $\hat{\underline{\theta}}_1^{EB(2)}$ which dominate $\underline{X}_1$ under the L_Q loss. They are defined as following.

$$\hat{\underline{\theta}}_1^{EB(1)}(\underline{X}) = (I_p - r(\underline{V}, \underline{S})\underline{V}\,\underline{S}^{-1})\underline{X}_1 + r(\underline{V}, \underline{S})\underline{V}\,\underline{S}^{-1}\underline{Z}, \tag{4.2.12}$$

where $r(\underline{V}, \underline{S})$ (> 0) is a scalar, and

$$\hat{\underline{\theta}}_1^{EB(2)}(\underline{X}) = (I_p - \underline{V}\,\underline{S}^{-1}\underline{g}(\underline{S}))\underline{X}_1 + \underline{V}\,\underline{S}^{-1}\underline{g}(\underline{S})\underline{Z}, \tag{4.2.13}$$

where $\underline{g}(\underline{S})$ is a pxp symmetric p.d. matrix. Also let

$$\hat{\underline{\theta}}_1^{EB(3)}(\underline{X}) = (I_p - r(\underline{V}, \underline{S})\underline{V}\,\underline{S}^{-1} - \underline{V}\,\underline{S}^{-1}\underline{f}(\underline{S}))\underline{X}_1 + (r(\underline{V}, \underline{S})\underline{V}\,\underline{S}^{-1} + \underline{V}\,\underline{S}^{-1}\underline{f}(\underline{S}))\underline{Z}, \tag{4.2.14}$$

where $\underline{f}(\underline{S})$ is a pxp symmetric p.d. matrix. It can be shown that $\hat{\underline{\theta}}_1^{EB(3)}$ beats $\hat{\underline{\theta}}_1^{EB(1)}$ under L_Q loss and under certain conditions on $r(\underline{V}, \underline{S})$ and $\underline{f}(\underline{S})$. As in Section 2.3, both $\underline{g}(\underline{S})$ and $\underline{f}(\underline{S})$ are of the form $\underline{P}\,\underline{\Psi}(\underline{\lambda})\underline{P}^T$ where $\underline{\Psi}(\underline{\lambda})$ is a diagonal matrix with ith diagonal element equal to $\Psi_i(\underline{\lambda}) > 0$, $i = 1, \ldots, p$. $\underline{\lambda} = (\lambda_1, \ldots, \lambda_p)^T$ denotes the vector of eigen values of $\underline{S}$, and $\underline{P}$ is an orthogonal matrix with its column vectors equal to the orthonormal eigen vectors of $\underline{S}$ ($\underline{S} = \underline{P}\,\Delta\,\underline{P}^T$, $\Delta = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$).

Remark 4.2.1 Suppose there are H (> 2) groups of sample matrix normal means $\underline{X}_h = (\underline{X}_{h1}, \ldots, \underline{X}_{hK})$, h = 1, ..., H. It is desired to estimate $\underline{\theta}_1 = (\underline{\theta}_{11}, \ldots, \underline{\theta}_{1K})$ on the basis of

$\underline{X}_h$'s since they are suspected to have the same population matrix means $\underline{\theta}_1$. In this case, we define $\underline{Z}_i = H^{-1}\sum_{h=1}^{H}\underline{X}_{hi}$, $\underline{Y}_i = H(\underline{X}_{1i}-\underline{Z}_i)$ and $\underline{S} = [H(H-1)]^{-1}\sum_{i=1}^{K}\underline{Y}_i\underline{Y}_i^T$, where marginally $\underline{Z}_i \sim N_p(\underline{\mu}_i, H^{-1}\Sigma)$, $\underline{Y}_i \sim N_p(\underline{0}, H(H-1)\Sigma)$ and $\underline{S} \sim W(p, K, \Sigma)$. It can be shown that the difference of risks of $\hat{\underline{\theta}}_1^{EB}$ and $\underline{X}_1$ is

$$
\begin{aligned}
&(H-1)H^{-1}E_{\underline{\theta}}tr\{[\underline{S}^{-1}\underline{\tau}(\underline{Y}, \underline{S})\underline{S}\,\underline{\tau}(\underline{Y}, \underline{S})\underline{S}^{-1} - 4\underline{D}^{S}[\underline{S}\,\underline{\tau}(\underline{Y}, \underline{S})\underline{S}^{-1}]\\
&-2(K-p-1)\underline{\tau}(\underline{Y}, \underline{S})\underline{S}^{-1}]\underline{Q}_0\} \qquad (4.2.15)
\end{aligned}
$$

The multiplier $(H-1)H^{-1}$ does not affect the way of finding EB minimax estimators $\hat{\underline{\theta}}_1^{EB(1)}$, $\hat{\underline{\theta}}_1^{EB(2)}$ and $\hat{\underline{\theta}}_1^{EB(3)}$ defined in (4.2.12)-(4.2.14), respectively.

4.3 Improvement On PTE

In this section we show how a modified EB estimator can dominate the PTE under the L_Q loss. A general PTE of $\underline{\theta}_1$ is given by

$$
\begin{aligned}
\hat{\underline{\theta}}_1^{PTE}(\underline{X}) &= g(C)\underline{X}_1+(1-g(C))\underline{Z}\\
&= [\underline{I}_p-(1-g(C))\underline{I}_p]\underline{X}_1+(1-g(C))\underline{Z}, \qquad (4.3.1)
\end{aligned}
$$

where $g(C) = 1_{[C>d]}$, $C = 2^{-1}\sum_{i=1}^{K}\underline{Y}_i^T\underline{V}^{-1}\underline{Y}_i = tr(\underline{V}^{-1}\underline{S})$, d being a positive constant depending on the chosen level of significance, and 1 the usual indicator function. Note that under the hypothesis of $\underline{\theta}_1 = \underline{\theta}_2$, C is distributed as χ^2_{Kp}. We propose the modified EB estimator

$$
\begin{aligned}
\hat{\underline{\theta}}_1^{MEB}(\underline{X}) &= g(C)\hat{\underline{\theta}}_1^{EB}+(1-g(C))\underline{Z}\\
&= [\underline{I}_p-[(1-g(C))\underline{I}_p+g(C)\underline{V}\,\underline{S}^{-1}\underline{\tau}_0(\underline{Y}, \underline{S})]]\underline{X}_1
\end{aligned}
$$

$$+[(1-g(C))\underline{I}_p+g(C)\underline{V}\,\underline{S}^{-1}\underline{\tau}_0(\underline{V},\underline{S})]\underline{Z} \qquad (4.3.2)$$

where $\underline{\tau}_0(\underline{V},\underline{S})$ is a p×p symmetric p.d. matrix. Note that $\widehat{\underline{\theta}}_1^{MEB}$ estimates $\underline{\theta}_1$ by $\underline{Z}$ if $g(C) = 0$, that is, the null hypothesis is accepted at a desired level of significance, and it estimates $\underline{\theta}_1$ by an EB estimator $\widehat{\underline{\theta}}_1^{EB}$ when the opposite is true. Also, $\widehat{\underline{\theta}}_1^{MEB}$ continues to be a weighted average of the pooled sample mean and the first sample mean which reflects the incorporation of degree of evidence for or against the null hypothesis. The following theorem is proved.

Theorem 4.3.1 Consider the L_Q loss, the difference of risks between $\widehat{\underline{\theta}}_1^{MEB}$ and $\widehat{\underline{\theta}}_1^{PTE}$ is given by

$$\begin{aligned}&R_Q(\underline{\theta}_1, \widehat{\underline{\theta}}_1^{MEB})-R_Q(\underline{\theta}_1, \widehat{\underline{\theta}}_1^{PTE})\\ &= \tfrac{1}{2}E_{\underline{\theta}}[g(C)[\underline{S}^{-1}\underline{\tau}_0(\underline{V},\underline{S})\underline{S}\,\underline{\tau}_0(\underline{V},\underline{S})\underline{S}^{-1}-4\underline{D}^S[\underline{S}\,\underline{\tau}_0(\underline{V},\underline{S})\underline{S}^{-1}]\\ &\quad -2(K-p-1)\underline{\tau}_0(\underline{V},\underline{S})\underline{S}^{-1}]\underline{Q}_0] \qquad (4.3.3)\end{aligned}$$

Proof of Theorem 4.3.1 Write $\underline{\tau}_1(\underline{V},\underline{S}) = (1-g(C))\underline{S}\,\underline{V}^{-1}$ and $\underline{\tau}_2(\underline{V},\underline{S}) = (1-g(C))\underline{S}\,\underline{V}^{-1} + g(C)\underline{\tau}_0(\underline{V},\underline{S}) = \underline{\tau}_1(\underline{V},\underline{S})+g(C)\underline{\tau}_0(\underline{V},\underline{S})$. Then $\widehat{\underline{\theta}}_1^{PTE}(\underline{X}) = (\underline{I}_p-\underline{V}\,\underline{S}^{-1}\underline{\tau}_1(\underline{V},\underline{S}))\underline{X}_1+\underline{V}\,\underline{S}^{-1}\underline{\tau}_1(\underline{V},\underline{S})\underline{Z}$ while $\widehat{\underline{\theta}}_1^{MEB}(\underline{X}) = (\underline{I}_p-\underline{V}\,\underline{S}^{-1}\underline{\tau}_2(\underline{V},\underline{S}))\underline{X}_1+\underline{V}\,\underline{S}^{-1}\underline{\tau}_2(\underline{V},\underline{S})\underline{Z}$. Note $\underline{\tau}_1(\underline{V},\underline{S})$ and $\underline{\tau}_2(\underline{V},\underline{S})$ are not symmetric. Let $\widehat{\underline{\theta}}_{11}^{MEB}(\underline{X}) = (\underline{I}_p-\underline{V}\,\underline{S}^{-1}\underline{\tau}_2(\underline{V},\underline{S}))\underline{X}_{11}+\underline{V}\,\underline{S}^{-1}\underline{\tau}_2(\underline{V},\underline{S})\underline{Z}_1$ and $\widehat{\underline{\theta}}_{11}^{PTE}(\underline{X}) = (\underline{I}_p-\underline{V}\,\underline{S}^{-1}\underline{\tau}_1(\underline{V},\underline{S}))\underline{X}_{11}+\underline{V}\,\underline{S}^{-1}\underline{\tau}_1(\underline{V},\underline{S})\underline{Z}_1$. Then, applying (4.2.2)-(4.2.4) twice, once with $\underline{\tau}(\underline{V},\underline{S}) = \underline{\tau}_1(\underline{V},\underline{S})$ and next with $\underline{\tau}(\underline{V},\underline{S}) = \underline{\tau}_2(\underline{V},\underline{S})$, one gets

$$\begin{aligned}&EL_Q(\underline{\theta}_1, \widehat{\underline{\theta}}_1^{MEB})-EL_Q(\underline{\theta}_1, \widehat{\underline{\theta}}_1^{PTE})\\ &= \tfrac{1}{2}E\mathrm{tr}\{(\underline{S}^{-1}\underline{\tau}_2(\underline{V},\underline{S})-\underline{\Sigma}^{-1})\underline{S}(\underline{\tau}_2^T(\underline{V},\underline{S})\underline{S}^{-1}-\underline{\Sigma}^{-1})\end{aligned}$$

$$(\underline{S}^{-1}\tau_1(\underline{Y},\underline{S})\underline{\Sigma}^{-1})\underline{S}(\tau_1^T(\underline{Y},\underline{S})\underline{S}^{-1}-\underline{\Sigma}^{-1})]\underline{Q}_0\} \qquad (4.3.4)$$

Since $\tau_2(\underline{Y},\underline{S}) = \tau_1(\underline{Y},\underline{S})+g(C)\tau_0(\underline{Y},\underline{S})$ and $\tau_0(\underline{Y},\underline{S})$ is symmetric, we rewrite the right-hand side of (4.3.4) as

$$\tfrac{1}{2}\,\mathrm{Etr}\{[g^2(C)\underline{S}^{-1}\tau_0(\underline{Y},\underline{S})\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}+g(C)\underline{S}^{-1}\tau_1(\underline{Y},\underline{S})\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}$$

$$+g(C)\underline{S}^{-1}\tau_0(\underline{Y},\underline{S})\underline{S}\,\tau_1^T(\underline{Y},\underline{S})\underline{S}^{-1}-g(C)\underline{S}^{-1}\tau_0(\underline{Y},\underline{S})\underline{S}\,\underline{\Sigma}^{-1}$$

$$-g(C)\underline{\Sigma}^{-1}\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}]\underline{Q}_0\} \qquad (4.3.5)$$

Also $g(C)\tau_1(\underline{Y},\underline{S}) = g(C)(1-g(C))\underline{S}\,\underline{Y}^{-1} = \underline{0}$, $g^2(C) = g(C)$, and $\mathrm{tr}(\underline{S}^{-1}\tau_0(\underline{Y},\underline{S})\underline{S}\,\underline{\Sigma}^{-1}\underline{Q}_0) = \mathrm{tr}(\underline{\Sigma}^{-1}\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}\underline{Q}_0)$. Hence,

$$EL_{\underline{Q}}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{MEB})-EL_{\underline{Q}}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{PTE})$$

$$= \tfrac{1}{2}\,\mathrm{Etr}\{[g(C)\underline{S}^{-1}\tau_0(\underline{Y},\underline{S})\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}-2\underline{\Sigma}^{-1}g(C)\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}]\underline{Q}_0\} \qquad (4.3.6)$$

Since marginally $S \sim W(p, K, \underline{\Sigma})$, apply the Wishart identity as given in Lemma 1.3.3 to get

$$E\{\underline{\Sigma}^{-1}[g(C)\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}]\}$$

$$= E\{2\underline{D}^{\underline{S}}[g(C)\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}]+(K-p-1)\underline{S}^{-1}[g(C)\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}]\}$$

$$= E\{2\underline{D}^{\underline{S}}g(C))\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}+2[(g(C)\underline{I}_p)^T(\underline{D}^{\underline{S}})^T]^T[\underline{S}\,\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}]$$

$$+(K-p-1)[g(C)\tau_0(\underline{Y},\underline{S})\underline{S}^{-1}]\}, \qquad (4.3.7)$$

where $\underline{D}^{\underline{S}}g(C) = ((\partial g(C)/\partial s_{ij}))$, $\partial g(C)/\partial s_{ij} = 0$ a.e. Thus from (4.3.6) and (4.3.7),

$$EL_{\underline{Q}}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{MEB})-EL_{\underline{Q}}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{PTE})$$

$$= \tilde{E}[R_Q(\underline{\theta}_1, \hat{\underline{\theta}}_1^{MEB}) - R_Q(\underline{\theta}_1, \hat{\underline{\theta}}_1^{PTE})]$$

$$= \tfrac{1}{2}\tilde{E} E_{\underline{\theta}}\{g(C)[\underline{S}^{-1} r_0(\underline{V}, \underline{S})\underline{S}\, r_0(\underline{V}, \underline{S})\underline{S}^{-1} - 4\underline{D}^{S}[\underline{S}\, r_0(\underline{V}, \underline{S})\underline{S}^{-1}]$$

$$- 2(K-p-1) r_0(\underline{V}, \underline{S})\underline{S}^{-1}]\underline{Q}_0\} \tag{4.3.8}$$

Again by using the completeness argument used in the proof of Theorem 4.2.1, (4.3.3) is the direct consequence of (4.3.8). The proof of the theorem is complete.

The above theorem provides an unbiased estimator of the risk difference of a modified EB estimator and the PTE, which can be recognized as g(C) times the unbiased estimator of the risk difference of the corresponding EB estimator and $\underline{X}_1$ by examining Theorem 4.2.1. Note that $\Pr(g(C) > 0) = \Pr(g(C) = 1) > 0$. Hence the EB minimax estimators $\hat{\underline{\theta}}_1^{EB(1)}$, $\hat{\underline{\theta}}_1^{EB(2)}$ and $\hat{\underline{\theta}}_1^{EB(3)}$ as given in (4.2.12), (4.2.13) and (4.2.14), can be modified to get $\hat{\underline{\theta}}_1^{MEB(1)}$, $\hat{\underline{\theta}}_1^{MEB(2)}$ and $\hat{\underline{\theta}}_1^{MEB(3)}$ respectively and they all dominate the PTE under the L_Q loss and under the minimaxity conditions. Note that

$$\hat{\underline{\theta}}_1^{MEB(1)}(\underline{X})$$

$$= [\underline{I}_p - \{(1-g(C))\underline{I}_p + g(C) r(\underline{V}, \underline{S})\underline{V}\,\underline{S}^{-1}\}]\underline{X}_1 + [(1-g(C))\underline{I}_p + g(C) r(\underline{V}, \underline{S})\underline{V}\,\underline{S}^{-1}]\underline{Z} \tag{4.3.9}$$

$$\hat{\underline{\theta}}_1^{MEB(2)}(\underline{X})$$

$$= [\underline{I}_p - \{(1-g(C))\underline{I}_p + g(C)\underline{V}\,\underline{S}^{-1}\underline{g}(\underline{S})\}]\underline{X}_1 + [(1-g(C))\underline{I}_p + g(C)\underline{V}\,\underline{S}^{-1}\underline{g}(\underline{S})]\underline{Z} \tag{4.3.10}$$

and

$$\hat{\underline{\theta}}_1^{MEB(3)}(\underline{X})$$

$$= \{\underline{I}_p - [(1-g(C))\underline{I}_p + g(C) r(\underline{V}, \underline{S})\underline{V}\,\underline{S}^{-1} + g(C)\underline{V}\,\underline{S}^{-1}\underline{f}(\underline{S})]\}\underline{X}_1$$

$$+[(1-g(C))\underline{I}_p+g(C)r(\underline{V},\underline{S})\underline{V}\,\underline{S}^{-1}+g(C)\underline{V}\,\underline{S}^{-1}f(\underline{S})]\underline{Z}. \tag{4.3.11}$$

CHAPTER 5
EMPIRICAL BAYES MINIMAX ESTIMATORS OF MATRIX NORMAL MEANS FOR ARBITRARY QUADRATIC LOSS AND UNKNOWN COVARIANCE MATRIX IN THE TWO-SAMPLE PROBLEM

5.1 Introduction

In Chapter 4, based on the assumption of a known covariance matrix $\underline{V}$, we investigated the estimation of one of the two sets of matrix normal means when it is suspected that the two matrix normal means are equal. In many practical situations, $\underline{V}$ is often unknown and needs to estimated from the supplementary data. Suppose $\hat{\underline{V}}$, a symmetric matrix independent of $\underline{X} = (\underline{X}_1, \underline{X}_2)$ is available as an estimate of $\underline{V}$. This is the situation in most multivariate normal models. Hence we continue the set up as (4.1.1) and (4.1.3), but relax the assumption of known $\underline{V}$ to the following three assumptions about $\underline{V}$ with a complete sufficient statistic $\hat{\underline{V}}$ independent of $\underline{X} = (\underline{X}_1, \underline{X}_2)$

(i) $\underline{V} = v\underline{G}$, $\underline{G}$ known, symmetric and p.d., and $v > 0$ unknown

$$\hat{\underline{V}} = \hat{v}\underline{G},\ \hat{v} \sim v(n+2)^{-1}\chi_n^2 \qquad (5.1.1)$$

where χ_n^2 is chi-square distribution with n degrees of freedom;

(ii) $\underline{V} = \underline{V}_d = \mathrm{diag}(v_1, \ldots, v_p)$, a diagonal matrix with ith diagonal element $v_i > 0$ and all v_i's unknown

$$\hat{\underline{V}}_d = \mathrm{diag}(\hat{v}_1, \ldots, \hat{v}_p),\ \hat{v}_i \sim v_i(n+2)^{-1}\chi_n^2 \text{ independently} \qquad (5.1.2)$$

(iii) $\underline{V}$ symmetric p.d. and entirely unknown. With

$$\hat{\underline{V}} \sim (n+p+1)^{-1}W(p, n, \underline{V}),\ n > p+1 \qquad (5.1.3)$$

In such cases the estimators of $\underline{\theta}_1$ are compared on the basis of general quadratic loss

$$L_{Q^*}(\underline{\theta}_1, \underline{a}) = \sum_{i=1}^{K} (\underline{\theta}_{1i} - \underline{a}_i)^T Q^* (\underline{\theta}_{1i} - \underline{a}_i), \tag{5.1.4}$$

where $Q^* = \underline{V}^{-1/2} Q \underline{V}^{-1/2}$, $\underline{V} = \underline{V}^{1/2}\underline{V}^{1/2}$, and again Q is a known p×p p.d. matrix. The multiplier $\underline{V}^{-1/2}$ used in the loss function because that makes $\underline{X}_1$ a minimax estimator of $\underline{\theta}_1$ with the constant risk Ktr(Q) not depending on any unknown parameter.

Next, we illustrate the availability of a complete sufficient statistic $\hat{\underline{V}}$ from the supplementary data for estimating the unknown $\underline{V}$ in a balanced two-way MANOVA model given in (4.1.2). First, suppose $\underline{V}_0 = v\underline{G}$, only v is unknown and $\underline{G}$ is p.d. Let $\hat{\underline{V}} = \hat{v}\underline{G}$, where $\hat{v} = \{m(2Kp(m-1)+2)\}^{-1}\mathrm{tr}(\sum_{h=1}^{2}\sum_{i=1}^{K}\sum_{j=1}^{m}(\underline{X}_{hij}-\underline{X}_{hi})\underline{G}^{-1}(\underline{X}_{hij}-\underline{X}_{hi})^T)$. Recall $\underline{X}_{hi} = m^{-1}\sum_{j=1}^{m}\underline{X}_{hij}$. Thus, $\underline{X}_{hi}|\underline{\theta}_{hi} \sim N_p(\underline{\theta}_{hi}, m^{-1}v\underline{G})$ and $\hat{v} \sim v\{m(2Kp(m-1)+2)\}^{-1}\chi^2_{2Kp(m-1)}$. If $\underline{V}_0 = \underline{V}_d = \mathrm{diag}(v_1, \ldots, v_p)$, an unknown diagonal matrix with ith diagonal element $v_i > 0$, let $\hat{\underline{V}} = \hat{\underline{V}}_d = \mathrm{diag}(\hat{v}_1, \ldots, \hat{v}_p)$, where $\hat{v}_i$ is $\{m(2K(m-1)+2)\}^{-1}$ times the ith diagonal element of $(\sum_{h=1}^{2}\sum_{i=1}^{K}\sum_{j=1}^{m}(\underline{X}_{hij}-\underline{X}_{hi})(\underline{X}_{hij}-\underline{X}_{hi})^T)$. In this case, $\underline{X}_{hi}|\underline{\theta}_{hi} \sim N_p(\underline{\theta}_{hi}, m^{-1}\underline{V}_d)$ and $\hat{v}_i \sim v_i\{m(2K(m-1)+2)\}^{-1}\chi^2_{2K(m-1)}$ independently. Finally, if $\underline{V}_0$ is complete unknown, then define $\hat{\underline{V}} = \{m(2K(m-1)+p+1)\}^{-1}(\sum_{h=1}^{2}\sum_{i=1}^{K}\sum_{j=1}^{m}(\underline{X}_{hij}-\underline{X}_{hi})(\underline{X}_{hij}-\underline{X}_{hi})^T)$. Now, $\underline{X}_{hi}|\underline{\theta}_{hi} \sim N_p(\underline{\theta}_{hi}, m^{-1}\underline{V}_0)$ and $\hat{\underline{V}} \sim \{m(2K(m-1)+p+1)\}^{-1}W(p, 2K(m-1), \underline{V}_0)$. In each of above three cases, the minimal sufficient statistic for $(\underline{\theta}_1, \underline{\theta}_2, \underline{V})$ is $(\underline{X}_1, \underline{X}_2, \hat{\underline{V}})$, where $\underline{V} = m^{-1}\underline{V}_0$. Note that conditional on $\underline{\theta} = (\underline{\theta}_1, \underline{\theta}_2)$, $\underline{X}_{hi}$'s and $\hat{\underline{V}}$ are independent, and the distribution of $\hat{\underline{V}}$ does not depend on $\underline{\theta}_1$ and $\underline{\theta}_2$. Also marginally, $(\underline{Z}, \underline{S}, \hat{\underline{V}})$ is the complete sufficient statistic for $(\underline{\mu}, \underline{\Sigma}, \underline{V})$, where $\underline{Z} = (\underline{Z}_1, \ldots, \underline{Z}_K)$, $\underline{Z}_i = (\underline{X}_{1i}+\underline{X}_{2i})/2 \sim N_p(\underline{\mu}_i, 2^{-1}\underline{\Sigma})$, $\underline{Y}_i = 2(\underline{X}_{1i}-\underline{Z}_i) \sim N_p(\underline{0}, 2\underline{\Sigma})$, and $\underline{S} = 2^{-1}(\sum_{i=1}^{K}\underline{Y}_i\underline{Y}_i^T) \sim W(p, K, \underline{\Sigma})$. Hence, the UMVUE of $\underline{\mu}$ is $\underline{Z}$, while the best multiple estimator of $\underline{\Sigma}^{-1}$ and $\underline{V}$ are $(K-p-1)\underline{S}^{-1}$ and $\hat{\underline{V}}$, respectively. Substituting these estimators for $\underline{\mu}$, $\underline{\Sigma}$ and $\underline{V}$ in (4.1.6), one gets the EB estimator of $\underline{\theta}_1$ as

$$(\underline{I}_p-(K-p-1)\hat{\underline{V}}\underline{S}^{-1})\underline{X}_1+(K-p-1)\hat{\underline{V}}\underline{S}^{-1}\underline{Z}. \qquad (5.1.5)$$

More generally, we consider the class of EB estimators

$$\hat{\underline{\theta}}_1^{EB}(\underline{X}, \hat{\underline{V}}) = (\underline{I}_p - \hat{\underline{V}}\underline{S}^{-1}\underline{\tau}(\hat{\underline{V}}, \underline{S}))\underline{X}_1+\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}(\hat{\underline{V}}, \underline{S})\underline{Z}, \qquad (5.1.6)$$

where $\underline{\tau}(\hat{\underline{V}}, \underline{S})$ is a p×p symmetric matrix.

In Section 5.2, Theorem 5.2.1 connects the risks of $\hat{\underline{\theta}}_1^{EB}$ and $\underline{X}_1$, and it shows that the methods of finding EB estimators in Section 3.2.3.4 can be adapted here for finding EB estimators of $\underline{\theta}_1$ under three different assumptions about $\underline{V}$ defined in (5.1.1)-(5.1.3), respectively. Again the EB estimators proposed in Section 5.2 are weighted averages of the pooled sample mean and the first sample mean. The weights intend to reflect the degree of evidence for or against the null hypothesis. Unlike the PTE, which selects either the pooled sample mean or the first sample mean, the EB estimator achieves a compromise between the two. It is shown in Section 5.3 that corresponding to every EB minimax estimator there exists a corresponding modified EB minimax estimator which dominates the PTE. The modified EB minimax estimator like the PTE is dichotomous. Both these estimators equal the pooled sample mean when the null hypothesis is accepted at a desired level of significance. However, if the opposite is the case, while the PTE equals first sample mean, the modified EB minimax estimator equals the corresponding EB minimax estimator.

5.2 EB Minimax Estimator Under The L_Q* Loss

In the next two sections, we denote by $E_{\underline{\theta}, \underline{V}}$, the expectation conditional on $\underline{\theta} = (\underline{\theta}_1, \underline{\theta}_2)$ and $\underline{V}$, $\tilde{E}$ the expectation over the marginal distribution of $\underline{\theta}$, and E the expectation over the joint distribution of $\underline{X} = (\underline{X}_1, \underline{X}_2)$, $\underline{\theta}$ and $\underline{V}$. The following theorem provide the risk of $\widehat{\underline{\theta}}_1^{EB}(\hat{\underline{V}}, \underline{X})$ defined in (5.1.6) under the L_{Q^*} loss without specifying the assumption regarding the structure of the unknown $\underline{V}$. We only need the independence of $\hat{\underline{V}}$ and $\underline{X} = (\underline{X}_1, \underline{X}_2)$ to prove the result.

Theorem 5.2.1 The risk $R_{Q^*}(\underline{\theta}_1, \widehat{\underline{\theta}}_1^{EB})$ under the loss given in (5.1.4) is given by

$$\begin{aligned} & R_{Q^*}(\underline{\theta}_1, \widehat{\underline{\theta}}_1^{EB}) \\ = \; & R_{Q^*}(\underline{\theta}_1, \underline{X}_1) + \tfrac{1}{2} E_{\underline{\theta}, \underline{V}} \mathrm{tr}\{\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S})\underline{S}\,\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^*\hat{\underline{V}} \\ & -4\underline{D}^S[\underline{S}\,\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}]\hat{\underline{V}}\underline{Q}^*\underline{V} - 2(K-p-1)\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^*\underline{V}\}, \end{aligned} \tag{5.2.1}$$

where $R_{Q^*}(\underline{\theta}_1, \underline{X}_1) = K\mathrm{tr}(\underline{V}\,\underline{Q}^*) = K\mathrm{tr}(\underline{Q})$

Proof of Theorem 5.2.1 Let $\widehat{\underline{\theta}}_{1i}^{B}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{X}_{1i} + \underline{V}\,\underline{\Sigma}^{-1}\underline{\mu}_i$ and $\widehat{\underline{\theta}}_{1i}^{EB}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - \hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S}))\underline{X}_{1i} + \hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S})\underline{Z}_i$, $i = 1, \ldots, K$. Standard calculation yields

$$\begin{aligned} & EL_{Q^*}(\underline{\theta}_1, \widehat{\underline{\theta}}_1^{EB}) \\ = \; & EL_{Q^*}(\underline{\theta}_1, \widehat{\underline{\theta}}_1^{B}) + E\{\sum_{i=1}^{K}(\widehat{\underline{\theta}}_{1i}^{B} - \widehat{\underline{\theta}}_{1i}^{EB})^T\underline{Q}^*(\widehat{\underline{\theta}}_{1i}^{B} - \widehat{\underline{\theta}}_{1i}^{EB})\} \end{aligned} \tag{5.2.2}$$

From (4.1.4) and (4.1.6), it follows that

$$EL_{Q^*}(\underline{\theta}_1, \widehat{\underline{\theta}}_1^{B}) = K\mathrm{tr}\{(\underline{I}_p - \underline{V}\,\underline{\Sigma}^{-1})\underline{V}\,\underline{Q}^*\} \tag{5.2.3}$$

Also $\widehat{\underline{\theta}}_{1i}^{B} - \widehat{\underline{\theta}}_{1i}^{EB}$ can be written as

$$\begin{aligned} & (\hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S}) - \underline{V}\,\underline{\Sigma}^{-1})(\underline{X}_{1i} - \underline{Z}_i) - \underline{V}\,\underline{\Sigma}^{-1}(\underline{Z}_i - \underline{\mu}_i) \\ = \; & \tfrac{1}{2}(\hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S}) - \underline{V}\,\underline{\Sigma}^{-1})\underline{Y}_i - \underline{V}\,\underline{\Sigma}^{-1}(\underline{Z}_i - \underline{\mu}_i) \end{aligned} \tag{5.2.4}$$

Next using the independence of $\underline{Z}_i$ and $(\underline{Y}_i, \hat{\underline{V}})$ and the fact that marginally $\underline{Z}_i \sim N_p(\underline{\mu}_i, 2^{-1}\Sigma)$, one gets

$$
\begin{aligned}
& E\{\sum_{i=1}^{K}(\hat{\underline{\theta}}_{1i}^{B}-\hat{\underline{\theta}}_{1i}^{EB})^{T}Q^{*}(\hat{\underline{\theta}}_{1i}^{B}-\hat{\underline{\theta}}_{1i}^{EB})\} \\
= \; & E\{\sum_{i=1}^{K}(\underline{Z}_i-\underline{\mu}_i)^{T}\Sigma^{-1}\underline{V}Q^{*}\underline{V}\Sigma^{-1}(\underline{Z}_i-\underline{\mu}_i) \\
& +\frac{1}{4}\sum_{i=1}^{K}\underline{Y}_i^{T}(\hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}},\underline{S})-\underline{V}\Sigma^{-1})^{T}Q^{*}(\hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}},\underline{S})-\underline{V}\Sigma^{-1})\underline{Y}_i\} \\
= \; & E\mathrm{tr}[(\sum_{i=1}^{K}(\underline{Z}_i-\underline{\mu}_i)(\underline{Z}_i-\underline{\mu}_i)^{T})\Sigma^{-1}\underline{V}Q^{*}\underline{V}\Sigma^{-1} \\
& +\frac{1}{2}(\hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}},\underline{S})-\underline{V}\Sigma^{-1})(\frac{1}{2}\sum_{i=1}^{K}\underline{Y}_i\underline{Y}_i^{T})(\tau^{T}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}-\Sigma^{-1}\underline{V})Q^{*}] \\
= \; & \frac{K}{2}\mathrm{tr}(\Sigma^{-1}\underline{V}Q^{*}\underline{V})+\frac{1}{2}E\mathrm{tr}\{(\hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}},\underline{S})-\underline{V}\Sigma^{-1})\underline{S}(\tau^{T}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}-\Sigma^{-1}\underline{V})Q^{*}\} \quad (5.2.5)
\end{aligned}
$$

Now we find

$$
\begin{aligned}
& \frac{1}{2}E\mathrm{tr}\{(\hat{\underline{V}}\underline{S}^{-1}\tau(\hat{\underline{V}},\underline{S})-\underline{V}\Sigma^{-1})\underline{S}(\tau^{T}(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}-\Sigma^{-1}\underline{V})Q^{*}\} \\
= \; & \frac{1}{2}E\mathrm{tr}\{\underline{S}^{-1}\tau(\hat{\underline{V}},\underline{S})\underline{S}\tau(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}Q^{*}\hat{\underline{V}}-2\Sigma^{-1}\underline{S}\tau(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}Q^{*}\underline{V}\} \\
& +\frac{K}{2}\mathrm{tr}(\Sigma^{-1}\underline{V}Q^{*}\underline{V}), \qquad (5.2.6)
\end{aligned}
$$

by the fact $E(\underline{S}) = K\Sigma$ and using the same algebra to get (3.1.5).

From (5.2.3), (5.2.5) and (5.2.6) it follows that the right-hand side of (5.2.2) is

$$
K\mathrm{tr}(\underline{V}Q^{*})+\frac{1}{2}E\mathrm{tr}\{\underline{S}^{-1}\tau(\hat{\underline{V}},\underline{S})\underline{S}\tau(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}Q^{*}\hat{\underline{V}}-2\Sigma^{-1}\underline{S}\tau(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}Q^{*}\underline{V}\} \quad (5.2.7)
$$

Since $\hat{\underline{V}}$ is independent of $\underline{X} = (\underline{X}_1, \underline{X}_2)$ and $\underline{S}$, we apply the Wishart identity as given in Lemma 1.3.3 again to $E[\Sigma^{-1}\underline{S}\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}]$. Thus $EL_{Q^*}(\underline{\theta}_1, \hat{\underline{\theta}}_1^{EB})$ can be expressed as

$$\tilde{E}R_{\underline{Q}^*}(\underline{\theta}_1, \hat{\underline{\theta}}_1^{EB})$$

$$= \tilde{E}R_{\underline{Q}^*}(\underline{\theta}_1, \underline{X}_1) + \frac{1}{2}\tilde{E}E_{\underline{\theta},\underline{Y}} tr[\underline{S}^{-1}\tau(\hat{\underline{V}}, \underline{S})\underline{S}\,\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^*\hat{\underline{V}}$$

$$-4\underline{D}^S[\underline{S}\,\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}]\hat{\underline{V}}\underline{Q}^*\underline{V} - 2(K-p-1)\tau(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^*\underline{V}\}, \tag{5.2.8}$$

where $R_{\underline{Q}^*}(\underline{\theta}_1, \underline{X}_1) = K tr(\underline{V}\,\underline{Q}^*)$ for all $\underline{\theta}_1$. Now the conditional distribution of $\underline{S} = 2^{-1}\sum_{i=1}^{K}\underline{Y}_i\underline{Y}_i^T$ given $\underline{\theta}$ depends on $\underline{\theta}$ only through $2^{-1}\sum_{i=1}^{K}(\underline{\theta}_{1i}-\underline{\theta}_{2i})(\underline{\theta}_{1i}-\underline{\theta}_{2i})^T = \underline{\Theta}^*$(say) and $\hat{\underline{V}}$ is ancillary. Also the family of distributions of $\underline{\Theta}$ being Wishart $W(p, K, \underline{\Delta})$ is complete for $\underline{\Delta}$ (or $\underline{\Sigma}$). Thus, (5.2.8) holds even when the $\tilde{E}$ expectation is removed. The proof of the theorem is complete.

Next observe the analogy between (3.1.3) and (5.2.1). As in Chapter 3, we have, therefore, two classes of EB minimax estimators,

$$\hat{\underline{\theta}}_1^{EB(1)}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - \tau(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\underline{S}^{-1})\underline{X}_1 + \tau(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\underline{S}^{-1}\underline{Z}, \tag{5.2.9}$$

where $\tau(\hat{\underline{V}}, \underline{S})$ (> 0) is a scalar, and

$$\hat{\underline{\theta}}_1^{EB(2)}(\hat{\underline{V}}, \underline{X}) = (\underline{I}_p - \hat{\underline{V}}\underline{S}^{-1}\underline{g}(\underline{S}))\underline{X}_1 + \hat{\underline{V}}\underline{S}^{-1}\underline{g}(\underline{S})\underline{Z}, \tag{5.2.10}$$

where $\underline{g}(\underline{S})$ is a pxp symmetric p.d. matrix and is of the form $\underline{P}\,\underline{\Psi}(\underline{\lambda})\underline{P}^T$, where $\underline{\Psi}(\underline{\lambda})$ is a diagonal matrix with ith diagonal element equal to $\Psi_i(\underline{\lambda}) > 0$, $i = 1, \ldots, p$, $\underline{\lambda} = (\lambda_1, \ldots, \lambda_p)^T$ denotes the vector of eigen values of $\underline{S}$, and $\underline{P}$ is an orthogonal matrix with its column vectors equal to the orthonormal eigen vectors of $\underline{S}$, that is, $\underline{S} = \underline{P}\,\underline{\Delta}\,\underline{P}^T$ and $\underline{\Delta} = diag(\lambda_1, \ldots, \lambda_p)$.

Also $\hat{\underline{\theta}}_1^{EB(1)}$ can be improved by

$$\hat{\underline{\theta}}_1^{EB(3)}(\hat{\underline{V}}, \underline{X})$$

$$= (\underline{I}_p - r(\underline{\hat{V}}, \underline{S})\underline{\hat{V}}\,\underline{S}^{-1}\,\underline{\hat{V}}\,\underline{S}^{-1}\underline{f}(\underline{S}))\underline{X}_1 + (r(\underline{\hat{V}}, \underline{S})\underline{\hat{V}}\,\underline{S}^{-1} + \underline{\hat{V}}\,\underline{S}^{-1}\underline{f}(\underline{S}))\underline{Z}, \qquad (5.2.11)$$

where $\underline{f}(\underline{S})$ is also of the form $\underline{P}\,\underline{\Psi}(\underline{\lambda})\underline{P}^T$ defined as above. The dominance holds for certain choice of $r(\underline{\hat{V}}, \underline{S})$ and $\underline{\Psi}(\underline{\lambda})$ which vary with the assumptions on $\underline{V}$, as shown in Section 3.2.3.4.

Remark 5.2.1 We can extend the number of groups of sample matrix normal means to H (> 2) and make the same adjustments for $\underline{Z}_i$ and $\underline{S}$ shown in Remark 4.2.1. In this case, the difference of risks of $\underline{\hat{\theta}}_1^{EB}$ and $\underline{X}_1$ is

$$(H-1)H^{-1}E_{\underline{\theta},\underline{V}}\,tr\{\underline{S}^{-1}r(\underline{\hat{V}}, \underline{S})\underline{S}\,r(\underline{\hat{V}}, \underline{S})\underline{S}^{-1}\underline{\hat{V}}\,\underline{Q}^*\underline{\hat{V}} - 4D^{\underline{S}}[\underline{S}\,r(\underline{\hat{V}}, \underline{S})\underline{S}^{-1}]\underline{\hat{V}}\,\underline{Q}^*\underline{V}$$

$$-2(K-p-1)r(\underline{\hat{V}}, \underline{S})\underline{S}^{-1}\underline{\hat{V}}\,\underline{Q}^*\underline{V}\} \qquad (5.2.12)$$

5.3 Improvement On PTE

When the covariance matrix $\underline{V}$ of $\underline{X}_{hi}$'s is unknown, a general PTE is given by

$$\begin{aligned}\underline{\hat{\theta}}_1^{PTE}(\underline{\hat{V}}, \underline{X}) &= g(C^*)\underline{X}_1 + (1-g(C^*))\underline{Z} \\ &= (\underline{I}_p - (1-g(C^*))\underline{I}_p)\underline{X}_1 + (1-g(C^*))\underline{Z}, \qquad (5.3.1)\end{aligned}$$

where $g(C^*) = I_{[C^*>d]}$ for some positive constant d and I denotes the usual indicator function. The choice of d is governed by the level of significance that is used for testing the null hypothesis of no difference between group means, that is $H_0: \underline{\theta}_1 = \underline{\theta}_2$. The statistic $C^* = C^*(\underline{\hat{V}}, \underline{S})$ is a scalar function of $\underline{\hat{V}}$ and $\underline{S}$. For example, when $\underline{V}$ is complete unknown, it can be Wilk's Lambda $(|\underline{\hat{V}}|/|\underline{\hat{V}}+\underline{S}|)$, Lawley Hotelling Trace statistic $tr(\underline{\hat{V}}\,\underline{S}^{-1})$, largest eigen-value of $(\underline{\hat{V}}\,\underline{S}^{-1})$ etc. To incorporate the degree of evidence for or against the null hypothesis that PTE fails to do, we propose the modified version of EB estimators $\hat{\underline{\theta}}_1^{EB}$ defined in (5.1.6) as

$$\begin{aligned}\hat{\underline{\theta}}_1^{MEB}(\hat{\underline{V}}, \underline{X}) &= g(C^*)\hat{\underline{\theta}}_1^{EB}+(1-g(C^*))\underline{Z} \\ &= [\underline{I}_p-\{(1-g(C^*))\underline{I}_p+g(C^*)\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}_0(\hat{\underline{V}}, \underline{S})\}]\underline{X}_1 \\ &\quad +\{(1-g(C^*))\underline{I}_p+g(C^*)\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}_0(\hat{\underline{V}}, \underline{S})\}\underline{Z}, \qquad (5.3.2)\end{aligned}$$

where $\underline{\tau}_0(\hat{\underline{V}}, \underline{S})$ is a pxp symmetric p.d. matrix. Note that $\hat{\underline{\theta}}_1^{MEB}$ is a weighted average of $\underline{X}_1$ and $\underline{Z}$ with extremes $\underline{Z}$ and $\hat{\underline{\theta}}_1^{EB}$ when $g(C^*)=0$ and 1, respectively.

The relationship between the risks of a PTE $\hat{\underline{\theta}}_1^{PTE}$ and a modified EB estimator $\hat{\underline{\theta}}_1^{MEB}$ is shown in the next theorem.

Theorem 5.3.1 The risk difference of $\hat{\underline{\theta}}_1^{MEB}$ and $\hat{\underline{\theta}}_1^{PTE}$ under the L_{Q^*} loss is given by

$$\begin{aligned}&R_{Q^*}(\underline{\theta}_1, \hat{\underline{\theta}}_1^{MEB})-R_{Q^*}(\underline{\theta}_1, \hat{\underline{\theta}}_1^{PTE}) \\ &= \tfrac{1}{2}E_{\underline{\theta},\underline{V}}\,\mathrm{tr}\{g(C^*)[\underline{S}^{-1}\underline{\tau}_0(\hat{\underline{V}}, \underline{S})\underline{S}\,\underline{\tau}_0(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^*\hat{\underline{V}} \\ &\quad -4\underline{D}^S\{\underline{S}\,\underline{\tau}_0(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\}\hat{\underline{V}}\underline{Q}^*\underline{V}-2(K-p-1)\underline{\tau}_0(\hat{\underline{V}}, \underline{S})\underline{S}^{-1}\hat{\underline{V}}\underline{Q}^*\underline{V}]\} \qquad (5.3.3)\end{aligned}$$

Proof of Theorem 5.3.1 Let $\underline{\tau}_1(\hat{\underline{V}}, \underline{S})=(1-g(C^*))\underline{S}\hat{\underline{V}}^{-1}$ and $\underline{\tau}_2(\hat{\underline{V}}, \underline{S})=(1-g(C^*))\underline{S}\hat{\underline{V}}^{-1}+g(C^*)\underline{\tau}_0(\hat{\underline{V}}, \underline{S})=\underline{\tau}_1(\hat{\underline{V}}, \underline{S})+g(C^*)\underline{\tau}_0(\hat{\underline{V}}, \underline{S})$. Then we rewrite $\hat{\underline{\theta}}_1^{PTE}(\hat{\underline{V}}, \underline{X})=(\underline{I}_p-\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}_1(\hat{\underline{V}}, \underline{S}))\underline{X}_1+\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}_1(\hat{\underline{V}}, \underline{S})\underline{Z}$ and $\hat{\underline{\theta}}_1^{MEB}(\hat{\underline{V}}, \underline{X})=(\underline{I}_p-\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}_2(\hat{\underline{V}}, \underline{S}))\underline{X}_1+\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}_2(\hat{\underline{V}}, \underline{S})\underline{Z}$. Note that $\underline{\tau}_0(\hat{\underline{V}}, \underline{S})$ is symmetric but $\underline{\tau}_1(\hat{\underline{V}}, \underline{S})$ and $\underline{\tau}_2(\hat{\underline{V}}, \underline{S})$ are not. Then let $\hat{\underline{\theta}}_{1i}^{PTE}(\hat{\underline{V}}, \underline{X})=(\underline{I}_p-\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}_1(\hat{\underline{V}}, \underline{S}))\underline{X}_{1i}+\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}_1(\hat{\underline{V}}, \underline{S})\underline{Z}_i$ and $\hat{\underline{\theta}}_{1i}^{MEB}(\hat{\underline{V}}, \underline{X})=(\underline{I}_p-\hat{\underline{V}}\underline{S}^{-1}\underline{\tau}_2(\hat{\underline{V}},$

$\underline{S}))\underline{X}_{11}+\underline{\hat{V}}\,\underline{S}^{-1}\underline{\tau}_2(\underline{\hat{V}},\underline{S})\underline{Z}_1$. Then, applying (5.2.2) and (5.2.5) twice, once with $\underline{\tau}(\underline{\hat{V}},\underline{S}) = \underline{\tau}_1(\underline{\hat{V}},\underline{S})$, and next with $\underline{\tau}(\underline{\hat{V}},\underline{S}) = \underline{\tau}_2(\underline{\hat{V}},\underline{S})$, we have

$$
\begin{aligned}
& EL_{\underline{Q}^*}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{MEB})\; EL_{\underline{Q}^*}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{PTE}) \\
= \; & \tfrac{1}{2}E\,\mathrm{tr}\{[(\underline{\hat{V}}\,\underline{S}^{-1}\underline{\tau}_2(\underline{\hat{V}},\underline{S})-\underline{V}\,\underline{\Sigma}^{-1})\underline{S}(\underline{\tau}_2^T(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\underline{\hat{V}}-\underline{\Sigma}^{-1}\underline{V}) \\
& -(\underline{\hat{V}}\,\underline{S}^{-1}\underline{\tau}_1(\underline{\hat{V}},\underline{S})-\underline{V}\,\underline{\Sigma}^{-1})\underline{S}(\underline{\tau}_1^T(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\underline{\hat{V}}-\underline{\Sigma}^{-1}\underline{V})]\underline{Q}^*\} \\
= \; & \tfrac{1}{2}E\,\mathrm{tr}\{[g^2(C^*)\underline{\hat{V}}\,\underline{S}^{-1}\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}\,\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\underline{\hat{V}} \\
& +g(C^*)\underline{\hat{V}}\,\underline{S}^{-1}\underline{\tau}_1(\underline{\hat{V}},\underline{S})\underline{S}\,\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\underline{\hat{V}}+g(C^*)\underline{\hat{V}}\,\underline{S}^{-1}\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}\,\underline{\tau}_1^T(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\underline{\hat{V}} \\
& -g(C^*)\underline{\hat{V}}\,\underline{S}^{-1}\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}\,\underline{\Sigma}^{-1}\underline{V}-g(C^*)\underline{V}\,\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\underline{\hat{V}}]\underline{Q}^*\}, \qquad (5.3.4)
\end{aligned}
$$

since $\underline{\tau}_2(\underline{\hat{V}},\underline{S}) = \underline{\tau}_1(\underline{\hat{V}},\underline{S})+g(C^*)\underline{\tau}_0(\underline{\hat{V}},\underline{S})$.

Next using the facts, $g(C^*)(1-g(C^*))=0$, $g^2(C^*)=g(C^*)$, and $\mathrm{tr}(\underline{\hat{V}}\,\underline{S}^{-1}\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}\,\underline{\Sigma}^{-1}\underline{V}\,\underline{Q}^*) = \mathrm{tr}(\underline{V}\,\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\underline{\hat{V}}\,\underline{Q}^*) = \mathrm{tr}(\underline{\Sigma}^{-1}\underline{S}\,\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\underline{\hat{V}}\,\underline{Q}^*\underline{V})$, one gets

$$
\begin{aligned}
& EL_{\underline{Q}^*}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{MEB})\; EL_{\underline{Q}^*}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{PTE}) \\
= \; & \tfrac{1}{2}E\,\mathrm{tr}\{g(C^*)\underline{S}^{-1}\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}\,\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\underline{\hat{V}}\,\underline{Q}^*\underline{\hat{V}} \\
& +2\underline{\Sigma}^{-1}g(C^*)\underline{S}\,\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\underline{\hat{V}}\,\underline{Q}^*\underline{V}\}. \qquad (5.3.5)
\end{aligned}
$$

Recall $\underline{\hat{V}}$ is independent of $\underline{X}$ and $\underline{S}$. By using the Wishart identity as given in Lemma 1.3.3 along with $\partial g(C^*)/\partial s_{ij} = 0$ a.e. and $\underline{D}^{\underline{S}}\, g(C^*) = ((\partial g(C^*)/\partial s_{ij})) = \underline{0}$, the following can be obtained.

$$E\{\underline{\Sigma}^{-1}g(C^*)\underline{S}\,\underline{\tau}_0(\underline{\hat{V}},\underline{S})\underline{S}^{-1}\}$$

$$= E[2\underline{D}^S[g(C^*)\underline{S}\tau_0(\hat{\underline{V}},\underline{S})\underline{S}^{-1}]+(K-p-1)\underline{S}^{-1}[g(C^*)\underline{S}\tau_0(\hat{\underline{V}},\underline{S})\underline{S}^{-1}]]$$

$$= E[2[\underline{D}^S g(C^*)]\underline{S}\tau_0(\hat{\underline{V}},\underline{S})\underline{S}^{-1}+2[(g(C^*)\underline{I}_p)^T(\underline{D}^S)^T]^T[\underline{S}\tau_0(\hat{\underline{V}},\underline{S})\underline{S}^{-1})$$

$$+(K-p-1)g(C^*)\tau_0(\hat{\underline{V}},\underline{S})\underline{S}^{-1}]$$

$$= E[2g(C^*)\underline{D}^S[\underline{S}\tau_0(\hat{\underline{V}},\underline{S})\underline{S}^{-1}]+(K-p-1)g(C^*)\tau_0(\hat{\underline{V}},\underline{S})\underline{S}^{-1}) \qquad (5.3.6)$$

Combining (5.3.5) and (5.3.6), one gets

$$E L_{Q^*}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{MEB})-E L_{Q^*}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{PTE})$$

$$= \bar{E}(E_{\underline{\theta},\underline{V}}L_{Q^*}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{MEB})-E_{\underline{\theta},\underline{V}}L_{Q^*}(\underline{\theta}_1,\hat{\underline{\theta}}_1^{PTE}))$$

$$= \frac{1}{2}\bar{E}E_{\underline{\theta},\underline{V}}\mathrm{tr}[g(C^*)[\underline{S}^{-1}\tau_0(\hat{\underline{V}},\underline{S})\underline{S}\tau_0(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^*\hat{\underline{V}}$$

$$-4\underline{D}^S[\underline{S}\tau_0(\hat{\underline{V}},\underline{S})\underline{S}^{-1}]\hat{\underline{V}}\,\underline{Q}^*\underline{V}-2(K-p-1)\tau_0(\hat{\underline{V}},\underline{S})\underline{S}^{-1}\hat{\underline{V}}\,\underline{Q}^*)\underline{V}]] \qquad (5.3.7)$$

Then applying the completeness argument as in the proof of Theorem 5.2.1 and the ancillarity of $\hat{\underline{V}}$ one gets (5.3.3) from (5.3.7) directly. The proof is complete.

In view of Theorem 5.2.1 and Theorem 5.3.1, unbiased estimators of the risk differences of $\hat{\underline{\theta}}_1^{EB}$ and $\underline{X}_1$, and $\hat{\underline{\theta}}_1^{MEB}$ and $\hat{\underline{\theta}}_1^{PTE}$ are different only in a multiplier $g(C^*)$. Note that $Pr(g(C^*) > 0) = Pr(g(C^*) = 1) > 0$. Hence under the L_{Q^*} loss, given an EB minimax estimator $\hat{\underline{\theta}}_1^{EB}$ and a PTE $\hat{\underline{\theta}}_1^{PTE}$ of $\underline{\theta}_1$, we can always produce a modified EB estimator $\hat{\underline{\theta}}_1^{MEB}$ which dominates $\hat{\underline{\theta}}_1^{PTE}$ under the same conditions where an EB minimax estimator $\hat{\underline{\theta}}_1^{EB}$ dominates $\underline{X}_1$. Thus, all the modified EB estimators $\hat{\underline{\theta}}_1^{MEB(1)}$, $\hat{\underline{\theta}}_1^{MEB(2)}$ and $\hat{\underline{\theta}}_1^{MEB(3)}$ of EB minimax estimators $\hat{\underline{\theta}}_1^{EB(1)}$, $\hat{\underline{\theta}}_1^{EB(2)}$ and $\hat{\underline{\theta}}_1^{EB(3)}$ as given in (5.2.9)-(5.2.11) respectively dominate the PTE under the L_{Q^*} loss. Note that

$$\hat{\underline{\theta}}_1^{MEB(1)}(\hat{\underline{V}}, \underline{X})$$

$$= [\underline{I}_p - [(1-g(C^*))\underline{I}_p + g(C^*)r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\underline{S}^{-1}]]\underline{X}_1 + [(1-g(C^*))\underline{I}_p + g(C^*)r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\underline{S}^{-1}]\underline{Z},$$

$$(5.3.8)$$

$$\hat{\underline{\theta}}_1^{MEB(2)}(\hat{\underline{V}}, \underline{X})$$

$$= [\underline{I}_p - [(1-g(C^*))\underline{I}_p + g(C^*)\hat{\underline{V}}\underline{S}^{-1}\underline{\varphi}(\underline{S})]]\underline{X}_1 + [(1-g(C^*))\underline{I}_p + g(C^*)\hat{\underline{V}}\underline{S}^{-1}\underline{\varphi}(\underline{S})]\underline{Z}, \quad (5.3.9)$$

and

$$\hat{\underline{\theta}}_1^{MEB(3)}(\hat{\underline{V}}, \underline{X})$$

$$= [\underline{I}_p - \{(1-g(C^*))\underline{I}_p + g(C^*)r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\underline{S}^{-1} + g(C^*)\hat{\underline{V}}\underline{S}^{-1}\underline{f}(\underline{S})\}]\underline{X}_1$$

$$+ [(1-g(C^*))\underline{I}_p + g(C^*)r(\hat{\underline{V}}, \underline{S})\hat{\underline{V}}\underline{S}^{-1} + g(C^*)\hat{\underline{V}}\underline{S}^{-1}\underline{f}(\underline{S})]\underline{Z} \qquad (5.3.10)$$

CHAPTER 6
SUMMARY AND FUTURE RESEARCH

6.1 Summary

This study, as noted in the previous chapters demonstrates Stein's phenomenon in the problem of estimation of matrix normal means. The estimators dominating the sample mean are obtained by EB methods which use information from similar sources in constructing estimators in addition to the most directly available source of information, or borrow strength from the ensemble

The proposed EB estimators are compared in term of risks under the quadratic loss with a p.d. weight matrix It should be noted that we only concentrate on estimators do not depend on the weight matrix being used. In this way the proposed EB estimators enjoy a great degree of robustness

Four usual assumptions about the covariance matrix are discussed in Chapters 2 and 3 Combined with the quadratic loss being considered, these two chapters cover all the situations that have been investigated in literature. In each case, three broad classes of EB minimax estimators of matrix normal means are proposed.

In Chapters 4 and 5, the problem of achieving a compromise between the first sample mean and pooled sample mean is solved by the proposed three classes of modified EB minimax estimators for both known and unknown covariance matrices. Each modified EB minimax estimator is a weighted average of the first sample mean and the pooled sample mean, and dominates the PTE in terms of risks under the quadratic loss Such estimators are modified EB minimax estimators which have uniformly smaller risks than the first sample mean.

6.2. Future Research

There are many directions in which to pursue future research, staying in the problem of estimation of matrix normal means. First, instead of using the same weight matrix for comparing the performance of some estimator in each column of the matrix normal means, one can use different weight matrices for each of them. We can also extend the results to multivariate regression models, for example generalizing the work of Ghosh, Lee and Littell (1988). Also, we propose to undertake the study of EB minimax estimators in growth curve models. Finally, we want to find also HB minimax estimators for such problems.

BIBLIOGRAPHY

Baranchik, A. J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution. Ann. Statist., 41, 642-645

Berger, J. D., and Bock, M. E. (1976a) Combining independent normal mean estimation problems with unknown variances. Ann. Statist., 4, 642-648

Berger, J. O., and Bock, M. E. (1976b) Improved minimax estimators of normal mean vectors for certain types of covariance matrices. Statistical Decision Theory and Related Topic II. S. S. Gupta and D. S. Moore (eds.) Academic Press. New York.

Berger, J. O., Bock, M. E., Brown, L. D., Casella, G., and Gleser, L. (1977) Minimax estimation of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix. Ann. Statist. 5, 763-771.

Berger, J. O., and Dey, D. K. (1983) Combining coordinates in simultaneous estimation of normal means. J. Statist. Planning and Inference., 8, 143-160

Berger, J. O., and Haff, L. R. (1983) A class of minimax estimators of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix. Statistics and Decisions., 1, 105-129

Brown, L. D. (1975) Estimation with incompletely specified loss functions. J. Amer. Statist. Assoc. 70, 417-427

Dey, D. K., Ghosh, M., and Srinivasan, C. (1985) Estimation of matrix means and the precision matrix: a unified approach. Unpublished.

Efron, B., and Morris, C. (1972a) Limiting the risk of Bayes and empirical Bayes estimators part II: the empirical Bayes case. J. Amer. Statist. Assoc., 67, 130-139

Efron, B., and Morris, C. (1972b) Empirical Bayes on vector observations: an extension of Stein's method. Biometrika., 59, 335-347

Efron, B., and Morris, C. (1973a) Stein's estimation rule and its competitors-an empirical Bayes approach. J. Amer. Statist. Assoc., 68, 117-130

Efron, B., and Morris, C. (1973b) Combining possibly related estimation problems. J. R. Statist. Soc., B., 35, 379-421

Efron, B., and Morris, C. (1976a) Family of minimax estimators of the mean of a multivariate normal distribution. Ann. Statist., 4, 11-21

Efron, B., and Morris, C. (1976b) Multivariate empirical Bayes and estimation of covariance matrices. Ann. Statist., 4, 22-32.

Ghosh, M. (1989) On the Γ minimaxity of a class of shrinkage estimators. Tech. Report No. 339 Department of Statistics, University of Florida.

Ghosh, M., Lee, L. C., and Littell, R. C. (1988) Empirical and hierarchical Bayes estimation of the response function. Tech. Report No. 309 Department of Statistics, University of Florida.

Ghosh, M. and Sinha, B. K. (1988) Empirical and hierarchical Bayes competitors of preliminary test estimators in two sample problems. J. Mult. Anal., 27, 206-227

Ghosh, M. and Shieh, G. (1990) Empirical Bayes minimax estimators of matrix normal means. Tech. Report No. 353 Department of Statistics, University of Florida.

Gleser, L. J. (1979) Minimax estimation of normal mean vector when the covariance matrix is unknown. Ann. Statist., 7, 838-846

Gleser, L. J. (1986a) Minimax estimators of normal mean vector for arbitrary quadratic loss and unknown covariance matrix. Ann. Statist., 14, 1625-1633

Gleser, L. J. (1986b) New estimators for the mean vector of a normal distribution with unknown covariance matrix. Statistical Decision Theory and Related Topic IV. S. S. Gupta and J. O. Berger (eds.) Academic Press, New York

Haff, L. R. (1977) Minimax estimators for a multinormal precision matrix. J. Mult. Anal., 7, 374-385

Haff, L. R. (1979) Estimation of the inverse covariance matrix: random mixtures of the inverse Wishart matrix and the identity. Ann. Statist., 7, 1264-1276

Haff, L. R. (1981) Further identities for the Wishart distribution with applications in regression. Can. J. Statist., 9, 215-244

Haff, L. R. (1982) Identities for the inverse Wishart distribution with computational results in linear and quadratic discrimination. Sankhya., B, 44, 245-258

Haff, L. R. (1984) Solutions of the Euler Largrange equations for certain multivariate normal equations. Unpublished.

James, W., and Stein, C. (1961) Estimation with quadratic loss. Proc. Fourth Berkeley Symp. Math. Statist. Prob., 1, 361-379 Univ. of California Press

Shinozaki, N. (1977) Simultaneous estimation of the means of independent variables with unknown variances. Keio Math. Sem. Rep., 2, 75-79

Shinozaki, N. (1980) Estimation of a multivariate normal mean with a class of quadratic loss functions. J. Amer. Statist. Assoc., 75, 973-976

Stein, C. (1956) Inadmissibility of the usual estimators for the mean of a multivariate normal distribution. Proc. Third Berkeley Symp. Math. Statist. Prob., 1, 197-206 Univ. of California Press.

Stein, C. (1974) Estimation of the mean of a multivariate normal distribution In Proc. Prague Symp. Asymptotic Statist., 345-381

Zheng, Z. (1986). On estimation of matrix of normal mean. J Mult. Anal., 18, 70-82.

BIOGRAPHICAL SKETCH

Gwowen Shieh was born in Kaohsiung, Taiwan, the Republic of China, in 1960. He graduated with a Bachelor of Science in management science from the National Chiao Tung University, Taiwan, R.O.C., in 1983. In the fall of 1984, he came to the U.S.A. and began graduate studies in the Department of Statistics and Probability at the Michigan State University. In 1985, he transferred to the University of Florida to continue graduate studies, receiving a Master of Science in statistics from the Michigan State University in 1986.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Malay Ghosh

Malay Ghosh, Chairman
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Mark C. K. Yang
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Beverly L. Brechner
Professor of Mathematics

This dissertation was submitted to the Graduate Faculty of the Department of Statistics in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy

December, 1990

Dean, Graduate School